

Shahman-Shah, K
09/848781
Seq. IDs 1-4 w/
inters
Page 1

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compen Ltd.

OM protein - protein search, using sw model

Run on: November 18, 2002, 15:53:56 ; Search time 118.742 Seconds
(without alignments)
48.867 Million cell updates/sec

Title: US-09-848-781-1
Perfect score: 43
Sequence: 1 NMLSEVERE 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues
Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- Pending_Patents_AA_Main:*
- 1: /cgn2_6/ptodata/1/paa/PCTUS_COMB.pep:*
- 2: /cgn2_6/ptodata/1/paa/US06_COMB.pep:*
- 3: /cgn2_6/ptodata/1/paa/US07_COMB.pep:*
- 4: /cgn2_6/ptodata/1/paa/US08_COMB.pep:*
- 5: /cgn2_6/ptodata/1/paa/US081_COMB.pep:*
- 6: /cgn2_6/ptodata/1/paa/US082_COMB.pep:*
- 7: /cgn2_6/ptodata/1/paa/US083_COMB.pep:*
- 8: /cgn2_6/ptodata/1/paa/US084_COMB.pep:*
- 9: /cgn2_6/ptodata/1/paa/US085_COMB.pep:*
- 10: /cgn2_6/ptodata/1/paa/US086_COMB.pep:*
- 11: /cgn2_6/ptodata/1/paa/US087_COMB.pep:*
- 12: /cgn2_6/ptodata/1/paa/US088_COMB.pep:*
- 13: /cgn2_6/ptodata/1/paa/US089_COMB.pep:*
- 14: /cgn2_6/ptodata/1/paa/US090_COMB.pep:*
- 15: /cgn2_6/ptodata/1/paa/US091_COMB.pep:*
- 16: /cgn2_6/ptodata/1/paa/US092_COMB.pep:*
- 17: /cgn2_6/ptodata/1/paa/US093_COMB.pep:*
- 18: /cgn2_6/ptodata/1/paa/US094_COMB.pep:*
- 19: /cgn2_6/ptodata/1/paa/US095_COMB.pep:*
- 20: /cgn2_6/ptodata/1/paa/US096_COMB.pep:*
- 21: /cgn2_6/ptodata/1/paa/US097_COMB.pep:*
- 22: /cgn2_6/ptodata/1/paa/US098_COMB.pep:*
- 23: /cgn2_6/ptodata/1/paa/US099_COMB.pep:*
- 24: /cgn2_6/ptodata/1/paa/US100_COMB.pep:*
- 25: /cgn2_6/ptodata/1/paa/US101_COMB.pep:*
- 26: /cgn2_6/ptodata/1/paa/US102_COMB.pep:*
- 27: /cgn2_6/ptodata/1/paa/US60_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	43	100.0	9	US-09-848-781-1	Sequence 1, Appl
2	38	88.4	772	US-09-570-581A-1397	Sequence 1397, Ap
3	38	88.4	5068	US-60-167-217-22900	Sequence 22900, A
4	38	88.4	5068	US-60-173-464-18575	Sequence 18575, A
5	38	88.4	5107	US-09-614-150-22563	Sequence 22563, A
6	38	88.4	5107	US-60-191-637-22640	Sequence 22640, A

7	38	88.4	5107	27	US-60-191-681-17870	Sequence 17870, A
8	34	79.1	102	15	US-09-107-532-5577	Sequence 5577, Ap
9	34	79.1	102	15	US-09-107-532A-5577	Sequence 5577, Ap
10	33	76.7	82	20	US-09-614-450-1279	Sequence 1279, Ap
11	33	76.7	284	21	US-09-791-537-5064	Sequence 5064, Ap
12	33	76.7	312	16	US-09-252-991A-25338	Sequence 25338, A
13	33	76.7	441	19	US-09-513-996A-24819	Sequence 24819, A
14	32	74.4	71	20	US-09-614-450-1280	Sequence 1280, Ap
15	32	74.4	101	27	US-60-192-739-3420	Sequence 3420, Ap
16	32	74.4	101	27	US-60-194-243-2474	Sequence 2474, Ap
17	32	74.4	172	1	PCT-US99-07478-2	Sequence 2, Appl
18	32	74.4	172	21	US-09-791-537-75562	Sequence 75562, A
19	32	74.4	215	1	PCT-US01-08631-42162	Sequence 42162, A
20	32	74.4	222	20	US-09-614-150-42774	Sequence 42774, A
21	32	74.4	222	27	US-60-167-324-2091	Sequence 2091, Ap
22	32	74.4	222	27	US-60-173-386-1876	Sequence 1876, Ap
23	32	74.4	222	27	US-60-175-871-2095	Sequence 2095, Ap
24	32	74.4	222	27	US-60-184-775-1809	Sequence 1809, Ap
25	32	74.4	222	27	US-60-191-637-42414	Sequence 42414, A
26	32	74.4	222	27	US-60-191-700-2037	Sequence 2037, Ap
27	32	74.4	334	16	US-09-248-796-19635	Sequence 19635, A
28	32	74.4	334	27	US-60-096-409-19635	Sequence 19635, A
29	32	74.4	430	19	US-09-513-996A-24820	Sequence 24820, A
30	32	74.4	556	25	US-10-179-131-8271	Sequence 8271, Ap
31	32	74.4	578	21	US-09-791-537-120562	Sequence 120562, A
32	32	74.4	629	26	US-10-238-075-1466	Sequence 1466, Ap
33	32	74.4	692	1	PCT-US01-08631-40291	Sequence 40291, A
34	32	74.4	922	18	US-09-417-507-42714	Sequence 42714, A
35	32	74.4	933	24	US-10-031-915-31	Sequence 31, Appl
36	32	74.4	1798	1	PCT-US01-08631-40292	Sequence 40292, A
37	32	74.4	1798	27	US-60-324-631-2646	Sequence 2646, Ap
38	32	74.4	1843	27	US-60-324-631-2646	Sequence 982, Ap
39	32	74.4	4011	1	PCT-US01-08631-42168	Sequence 42168, A
40	32	74.4	7718	1	PCT-US01-08631-42170	Sequence 42170, A
41	31	72.1	200	21	US-09-791-537-147585	Sequence 147585, A
42	31	72.1	263	18	US-09-489-038A-8239	Sequence 8239, Ap
43	31	72.1	277	21	US-09-791-537-96220	Sequence 96220, A
44	31	72.1	283	20	US-09-614-150-5487	Sequence 5487, Ap
45	31	72.1	283	27	US-60-167-217-5601	Sequence 5601, Ap

ALIGNMENTS

RESULT 1
US-09-848-781-1
; Sequence 1, Application US/09848781
; GENERAL INFORMATION:
; APPLICANT: SANDERS, MITCHELL
; TITLE OF INVENTION: A DEVICE FOR DETECTING BACTERIAL CONTAMINATION AND METHOD OF U
; FILE REFERENCE: 102951-10
; CURRENT APPLICATION NUMBER: US/09/848,781
; PRIOR FILING DATE: 2001-05-03
; PRIOR APPLICATION NUMBER: 60/201,405
; PRIOR FILING DATE: 2000-05-03
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Listeria monocytogenes
US-09-848-781-1

Query Match 100.0% ; Score 43 ; DB 22 ; Length 9 ;
Best Local Similarity 100.0% ; Pred No. 4.2e+06 ;
Matches 9 ; Conservative 0 ; Mismatches 0 ; Indels 0 ; Gaps 0 ;

QY 1 NMLSEVERE 9
|||||
Db 1 NMLSEVERE 9

RESULT 2

US-09-570-581A-1397
; Sequence 1397, Application US/09570581A
; GENERAL INFORMATION:
; APPLICANT: Ceres Inc.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-0875P
; CURRENT APPLICATION NUMBER: US/09/570,581A
; CURRENT FILING DATE: 2001-05-12
; NUMBER OF SEQ ID NOS: 1950
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1397
; LENGTH: 772
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-570-581A-1397

Query Match 88.4%; Score 38; DB 19; Length 772;
Best Local Similarity 77.8%; Pred. No. 1.5e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 NMLSEVERE 9
||| |||
Db 263 NMLDEIERE 271

RESULT 3
; Sequence 22900, Application US/60167217
; GENERAL INFORMATION:
; APPLICANT: Li, Peter W. D.
; TITLE OF INVENTION: ISOLATED DROSOPHILA PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING DROSOPHILA PROTEINS, AND USES
; FILE REFERENCE: CL000152
; CURRENT APPLICATION NUMBER: US/60/167,217
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 23195
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22900
; LENGTH: 5068
; TYPE: PRT
; ORGANISM: Drosophila
US-60-167-217-22900

Query Match 88.4%; Score 38; DB 27; Length 5068;
Best Local Similarity 88.9%; Pred. No. 1.5e+03;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 NMLSEVERE 9
||||| |||
Db 2867 NMLSEVERE 2875

RESULT 4
; Sequence 18575, Application US/60173464
; GENERAL INFORMATION:
; APPLICANT: Li, Peter W. D.
; TITLE OF INVENTION: ISOLATED G-PROTEIN COUPLED RECEPTORS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING GPCR PROTEINS AND USES
; FILE REFERENCE: CL000173
; CURRENT APPLICATION NUMBER: US/60/173,464
; CURRENT FILING DATE: 1999-12-29
; NUMBER OF SEQ ID NOS: 30269
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18575
; LENGTH: 5068
; TYPE: PRT
; ORGANISM: Drosophila
US-60-173-464-18575

Query Match 88.4%; Score 38; DB 27; Length 5068;
Best Local Similarity 88.9%; Pred. No. 1.5e+03;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 NMLSEVERE 9
||||| |||
Db 2867 NMLSEVERE 2875

RESULT 5
; Sequence 22563, Application US/09614150
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/09/614,150
; CURRENT FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22563
; LENGTH: 5107
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-09-614-150-22563

Query Match 88.4%; Score 38; DB 20; Length 5107;
Best Local Similarity 88.9%; Pred. No. 1.6e+03;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 NMLSEVERE 9
||||| |||
Db 2871 NMLSEVERE 2879

RESULT 6
; Sequence 22640, Application US/60191637
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; TITLE OF INVENTION: NUCLEIC ACID DETECTION KITS COMPRISING
; TITLE OF INVENTION: GENE SEQUENCES EXPRESSED FROM THE DROSOPHILA GENOME, AND
; FILE REFERENCE: CL000392
; CURRENT APPLICATION NUMBER: US/60/191,637
; CURRENT FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 42660
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22640
; LENGTH: 5107
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-60-191-637-22640

Query Match 88.4%; Score 38; DB 27; Length 5107;
Best Local Similarity 88.9%; Pred. No. 1.6e+03;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NMLSEVERE 9
|||||
Db 2871 NMLSEVERE 2879

RESULT 7
US-60-191-681-17870
; Sequence 17870, Application US/60191681
; GENERAL INFORMATION:
; APPLICANT: LI, Peter, W.D.
; TITLE OF INVENTION: ISOLATED DROSOPHILA PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING ISOLATED DROSOPHILA PROTEINS AND
; FILE REFERENCE: c1000390
; CURRENT APPLICATION NUMBER: US/60/191,681
; CURRENT FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 30973
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17870
; LENGTH: 5107
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-60-191-681-17870

Query Match 88.4%; Score 38; DB 27; Length 5107;
Best Local Similarity 88.9%; Pred. No. 1.6e+03;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NMLSEVERE 9
|||||
Db 2871 NMLSEVERE 2879

RESULT 8
US-09-107-532-5577
; Sequence 5577, Application US/09107532
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECIUM FOR
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7308
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER:
; OPERATING SYSTEM:
; SOFTWARE:
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085598
; FILING DATE: May 14, 1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007

TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 5577:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 102 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1...102
US-09-107-532-5577

Query Match 79.1%; Score 34; DB 15; Length 102;
Best Local Similarity 87.5%; Pred. No. 62;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 MLESEVERE 9
|||||
Db 73 MLESEVERE 80

RESULT 9
US-09-107-532A-5577
; Sequence 5577, Application US/09107532A
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 5577:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 102 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (B) LOCATION 1...102
; SEQUENCE DESCRIPTION: SEQ ID NO: 5577:
US-09-107-532A-5577

Query Match 79.1%; Score 34; DB 15; Length 102;
Best Local Similarity 87.5%; Pred. No. 62;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 MLSEVERE 9
Db 73 MLSELERE 80

RESULT 10

US-09-614-450-1279
; Sequence 1279, Application US/09614450
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Thereby
; FILE REFERENCE: 2750-1061P
; CURRENT APPLICATION NUMBER: US/09/614,450
; CURRENT FILING DATE: 2000-07-14
; NUMBER OF SEQ ID NOS: 5797
; SEQ ID NO 1279
; LENGTH: 82
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: 1..82
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc-feature
; LOCATION: 1..82
; OTHER INFORMATION: Ceres Seq. ID 1384435
US-09-614-450-1279

Query Match 76.7%; Score 33; DB 20; Length 82;
Best Local Similarity 66.7%; Pred. No. 77;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 NMLSEVERE 9
Db 11 SMLDEIERE 19

RESULT 11

US-09-791-537-5064
; Sequence 5064, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomimix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5064
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Streptomyces coelicolor A3
US-09-791-537-5064

Query Match 76.7%; Score 33; DB 21; Length 284;
Best Local Similarity 66.7%; Pred. No. 4e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 NMLSEVERE 9
Db 208 NMLSEVERE 216

RESULT 12

US-09-252-991A-25338
; Sequence 25338, Application US/09252991A
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25338
; LENGTH: 312
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25338

Query Match 76.7%; Score 33; DB 16; Length 312;
Best Local Similarity 87.5%; Pred. No. 4.5e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 MLSEVERE 9
Db 177 MLSEVERE 184

RESULT 13

US-09-513-996A-24819
; Sequence 24819, Application US/09513996A
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
; TITLE OF INVENTION: ENCODED THEREBY
; FILE REFERENCE: 2750-709P
; CURRENT APPLICATION NUMBER: US/09/513,996A
; CURRENT FILING DATE: 2000-02-25
; NUMBER OF SEQ ID NOS: 81028
; SEQ ID NO 24819
; LENGTH: 441
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 1..441
; OTHER INFORMATION: any n or Xaa = unknown
; FEATURE:
; OTHER INFORMATION: Location 1..441 / Ceres Seq. ID 1534067
US-09-513-996A-24819

Query Match 76.7%; Score 33; DB 19; Length 441;
Best Local Similarity 66.7%; Pred. No. 7.2e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 NMLSEVERE 9
Db 11 SMLDEIERE 19

RESULT 14

US-09-614-450-1280
; Sequence 1280, Application US/09614450
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptid
; FILE REFERENCE: 2750-1061P
; CURRENT APPLICATION NUMBER: US/09/614,450
; CURRENT FILING DATE: 2000-07-14
; NUMBER OF SEQ ID NOS: 5797
; SEQ ID NO 1280


```

; LENGTH: 71
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..71
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..71
; OTHER INFORMATION: Ceres Seq. ID 1384436
US-09-614-450-1280

```

```

Query Match          74.4%; Score 32; DB 20; Length 71;
Best Local Similarity 75.0%; Pred. No. 1e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      2 MISEVRE 9
        1111111
Db       1 MIDEIERE 8

```

```

RESULT 15
US-60-192-739-3420
; Sequence 3420, Application US/60192739
; GENERAL INFORMATION:
; APPLICANT: Bonazzi, Vivien
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: C1000406
; CURRENT APPLICATION NUMBER: US/60/192,739
; CURRENT FILING DATE: 2000-03-28
; NUMBER OF SEQ ID NOS: 4532
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3420
; LENGTH: 101
; TYPE: PRT
; ORGANISM: HUMAN
US-60-192-739-3420

```

```

Query Match          74.4%; Score 32; DB 27; Length 101;
Best Local Similarity 75.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      2 MISEVRE 9
        1111111
Db       3 MISEVDRE 10

```

Search completed: November 18, 2002, 16:05:04
 Job time : 121.742 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 18, 2002, 15:55:36 ; Search time 3.19355 Seconds
(without alignments)
36.183 Million cell updates/sec

Title: US-09-848-781-1

Perfect score: 43

Sequence: 1 NMLSEVERE 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 47062 seqs, 12839170 residues

Total number of hits satisfying chosen parameters: 47062

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_AA_New:*

1: /cgn2_6/ptodata/2/paa/PC1_NEW_COMB.pep:.*
2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep:.*
3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep:.*
4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep:.*
5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep:.*
6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep:.*
7: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep:.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	29	67.4	61	6	US-10-259-520-9
2	29	67.4	242	1	PCT-US02-32852-12
3	29	67.4	645	6	US-10-120-801-101
4	29	67.4	664	6	US-10-120-801-100
5	28	65.1	92	6	US-10-092-411A-4014
6	28	65.1	117	6	US-10-092-411A-4706
7	28	65.1	172	5	US-09-513-999C-7944
8	28	65.1	173	5	US-09-513-999C-5427
9	28	65.1	198	5	US-09-513-999C-5428
10	28	65.1	244	6	US-10-198-070-46
11	28	65.1	344	6	US-10-092-411A-3158
12	28	65.1	429	1	PCT-US02-32032-16
13	28	65.1	429	6	US-10-131-813A-94
14	28	65.1	429	6	US-10-131-819A-94
15	28	65.1	429	6	US-10-131-823A-94
16	28	65.1	429	6	US-10-131-824A-94
17	28	65.1	429	6	US-10-131-826A-94
18	28	65.1	429	6	US-10-131-829A-94
19	28	65.1	429	6	US-10-125-926A-94
20	28	65.1	429	6	US-10-127-829A-94
21	28	65.1	429	6	US-10-127-831A-94
22	28	65.1	429	6	US-10-127-835A-94
23	28	65.1	429	6	US-10-127-837A-94
24	28	65.1	429	6	US-10-127-842A-94
25	28	65.1	429	6	US-10-127-850A-94
26	28	65.1	429	6	US-10-127-901A-94

27	28	65.1	429	6	US-10-128-689A-94	Sequence 94, Appl
28	28	65.1	429	6	US-10-131-830A-94	Sequence 94, Appl
29	28	65.1	429	6	US-10-131-833A-94	Sequence 94, Appl
30	28	65.1	429	6	US-10-131-837A-94	Sequence 94, Appl
31	28	65.1	429	6	US-10-125-930A-94	Sequence 94, Appl
32	28	65.1	429	6	US-10-127-825A-94	Sequence 94, Appl
33	28	65.1	429	6	US-10-127-838B-94	Sequence 94, Appl
34	28	65.1	429	6	US-10-127-843A-94	Sequence 94, Appl
35	28	65.1	429	6	US-10-128-684A-94	Sequence 94, Appl
36	28	65.1	429	6	US-10-128-685A-94	Sequence 94, Appl
37	28	65.1	429	6	US-10-128-686A-94	Sequence 94, Appl
38	28	65.1	429	6	US-10-128-690A-94	Sequence 94, Appl
39	28	65.1	429	6	US-10-128-693A-94	Sequence 94, Appl
40	28	65.1	429	6	US-10-131-821A-94	Sequence 94, Appl
41	28	65.1	429	6	US-10-131-836A-94	Sequence 94, Appl
42	28	65.1	429	6	US-10-137-872A-94	Sequence 94, Appl
43	28	65.1	429	6	US-10-137-873A-94	Sequence 94, Appl
44	28	65.1	429	6	US-10-125-921A-94	Sequence 94, Appl
45	28	65.1	429	6	US-10-125-921A-94	Sequence 94, Appl

ALIGNMENTS

RESULT 1
US-10-259-520-9
; Sequence 9, Application US/10259520
GENERAL INFORMATION:
APPLICANT: Jian Ni et al.
TITLE OF INVENTION: Mammary Transforming Protein
NUMBER OF SEQUENCE: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: HUMAN GENOME SCIENCES, INC
STREET: 9410 KEY WEST AVENUE
CITY: ROCKVILLE
STATE: MARYLAND
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 INCH DISKETTE
OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/259,520
FILING DATE: 30-Sep-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/263,811
FILING DATE: MAR-08-1999
APPLICATION NUMBER: 08/743,975
FILING DATE: 01-NOV-1996
APPLICATION NUMBER: 60/006,187
FILING DATE: 02 NOVEMBER 1995
ATTORNEY/AGENT INFORMATION:
NAME: MICHELE M. WALES
REGISTRATION NUMBER: 43,975
REFERENCE/DOCKET NUMBER: PF212D1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 301-309-8504
TELEFAX: 301-309-8439
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 61 AMINO ACIDS
TYPE: AMINO ACID
STRANDEDNESS: <Unknown>
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
SEQUENCE DESCRIPTION: SEQ ID NO: 9:
US-10-259-520-9
Query Match 67.4%; Score 29; DB 6; Length 61;
Best Local Similarity 66.7%; Pred. No. 7.9;

Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 NMLEVERE 9
|:|:|
Db 19 NLSRPERE 27

RESULT 2
PCT-US02-32852-12
; Sequence 12, Application PC/TUS0232852
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: TANG, Y. Tom
; APPLICANT: FORSYTHE, Ian J.
; APPLICANT: EMERLING, Brooke M.
; APPLICANT: HAFALIA, April J.A.
; APPLICANT: YUE, Henry
; APPLICANT: XU, Yuming
; APPLICANT: GIETZEN, Kimberly J.
; APPLICANT: CHAWLA, Nalinder K.
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: MARQUIS, Joseph P.
; APPLICANT: BECHA, Shanya D.
; APPLICANT: KABLE, Amy E.
; APPLICANT: LAL, Preeti G.
; APPLICANT: RICHARDSON, Thomas W.
; APPLICANT: LEE, Soo Yeun
; APPLICANT: LEE, Ernestine A.
; APPLICANT: TRAN, Bao
; APPLICANT: WARREN, Bridget A.
; APPLICANT: LU, Dzung Anna M.
; APPLICANT: GURURAJAN, Rajagopal
; APPLICANT: SPRAGUE, William W.
; APPLICANT: BLAKE, Julie J.
; APPLICANT: THANCAVELU, Kavitha
; APPLICANT: SMARAKAR, Anita
; APPLICANT: CORVAD, Ann E.
; APPLICANT: GRIFFIN, Jennifer A.
; APPLICANT: LINDQUIST, Erika A.
; APPLICANT: ELLIOTT, Vicki S.
; APPLICANT: ISON, Craig H.
; APPLICANT: RAMKUMAR, Jayaxmi
; TITLE OF INVENTION: MOLECULES FOR DISEASE DETECTION AND TREATMENT
; FILE REFERENCE: PF-1232 PCT
; CURRENT APPLICATION NUMBER: PCT/US02/32852
; CURRENT FILING DATE: 2002-10-10
; PRIOR APPLICATION NUMBER: US 60/328,944
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/345,384
; PRIOR FILING DATE: 2001-10-26
; PRIOR APPLICATION NUMBER: US 60/343,880
; PRIOR FILING DATE: 2001-11-02
; PRIOR APPLICATION NUMBER: US 60/345,143
; PRIOR FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: US 60/332,430
; PRIOR FILING DATE: 2001-11-16
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PERL Program
; SEQ ID NO 12
; LENGTH: 242
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 7502593CD1
PCT-US02-32852-12

Query Match 67.4%; Score 29; DB 1; Length 242;
Best Local Similarity 62.5%; Pred. No. 35;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
QY 2 MLEVERE 9
|:|:|:|

Db 113 ILSEIERE 120

RESULT 3
US-10-120-801-101
; Sequence 101, Application US/10120801
; GENERAL INFORMATION:
; APPLICANT: Pena, Carol
; APPLICANT: Guo, Xiaojia
; APPLICANT: Shimkets, Richard
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Spytek, Kimberly
; APPLICANT: Mehraban, Fuad
; APPLICANT: Topper, James N.
; APPLICANT: Malysankar, Uriel
; APPLICANT: Wasserman, Scott
; APPLICANT: Edinger, Shlomit
; APPLICANT: Smithson, Glenda
; APPLICANT: Gunther, Erik
; APPLICANT: Komuves, Laszlo
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-340
; CURRENT APPLICATION NUMBER: US/10/120,801
; CURRENT FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: 60/285748
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: 60/286068
; PRIOR FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: 60/286292
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: 60/288334
; PRIOR FILING DATE: 2001-05-03
; PRIOR APPLICATION NUMBER: 60/291241
; PRIOR FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: 60/322284
; PRIOR FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: 60/285609
; PRIOR FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 101
; LENGTH: 645
; TYPE: PRT
; ORGANISM: mouse
US-10-120-801-101

Query Match 67.4%; Score 29; DB 6; Length 645;
Best Local Similarity 66.7%; Pred. No. 1e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 NMLEVERE 9
|:|:|:|
Db 96 NLSLEVSKE 104

RESULT 4
US-10-120-801-100
; Sequence 100, Application US/10120801
; GENERAL INFORMATION:
; APPLICANT: Pena, Carol
; APPLICANT: Guo, Xiaojia
; APPLICANT: Shimkets, Richard
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Spytek, Kimberly
; APPLICANT: Mehraban, Fuad
; APPLICANT: Topper, James N.
; APPLICANT: Malysankar, Uriel
; APPLICANT: Wasserman, Scott
; APPLICANT: Edinger, Shlomit
; APPLICANT: Smithson, Glenda
; APPLICANT: Gunther, Erik

```
APPLICANT: Komuves, Laszlo
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-340
; CURRENT APPLICATION NUMBER: US/10/120,801
; CURRENT FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: 60/285748
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: 60/286068
; PRIOR FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: 60/286292
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: 60/288334
; PRIOR FILING DATE: 2001-05-03
; PRIOR APPLICATION NUMBER: 60/291241
; PRIOR FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: 60/322284
; PRIOR FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: 60/285609
; PRIOR FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 100
; LENGTH: 664
; TYPE: PRT
; ORGANISM: mouse
US-10-120-801-100
```

```
Query Match          67.4%; Score 29; DB 6; Length 664;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
```

```
OY 1 NMLSEVERE 9
    1 111111
DB 115 NMLSEVERE 123
```

```
RESULT 5
US-10-092-411A-4014
; Sequence 4014, Application US/10092411A
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: 032796-101
; CURRENT APPLICATION NUMBER: US/10/092,411A
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 09/134,001
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5676
; SEQ ID NO 4014
; LENGTH: 92
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-10-092-411A-4014
```

```
Query Match          65.1%; Score 28; DB 6; Length 92;
Best Local Similarity 55.6%; Pred. No. 20;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
```

```
OY 1 NMLSEVERE 9
    1 111111
DB 67 NMLTEVEAD 75
```

```
RESULT 6
US-10-092-411A-4706
; Sequence 4706, Application US/10092411A
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
```

```
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-101
; CURRENT APPLICATION NUMBER: US/10/092,411A
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 09/134,001
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5676
; SEQ ID NO 4706
; LENGTH: 117
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-10-092-411A-4706
```

```
Query Match          65.1%; Score 28; DB 6; Length 117;
Best Local Similarity 55.6%; Pred. No. 26;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
```

```
OY 1 NMLSEVERE 9
    1 111111
DB 17 NMLTELENE 25
```

```
RESULT 7
US-09-513-999C-7944
; Sequence 7944, Application US/09513999C
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclet, A.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: 59,US2,REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 7944
; LENGTH: 172
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 101
; OTHER INFORMATION: Xaa=Asp or Asn
US-09-513-999C-7944
```

```
Query Match          65.1%; Score 28; DB 5; Length 172;
Best Local Similarity 50.0%; Pred. No. 39;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 2 MLESEVERE 9
    1 111111
DB 72 MLESEIDKE 79
```

```
RESULT 8
US-09-513-999C-5427
; Sequence 5427, Application US/09513999C
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclet, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: 59,US2,REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
```

;; PRIOR FILING DATE: 1999-02-26
;; NUMBER OF SEQ ID NOS: 36681
;; SOFTWARE: Patent.pm
;; SEQ ID NO 5427
;; LENGTH: 173
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-513-999C-5427

Query Match
Best Local Similarity 62.5%; Score 28; DB 5; Length 173;
Best Local Similarity 62.5%; Pred. No. 39;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NMLSEVER 8
Db 83 NMLSELEK 90

RESULT 9
US-09-513-999C-5428
; Sequence 5428, Application US/09513999C
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: 59. US2. REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 5428
; LENGTH: 198
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-513-999C-5428

Query Match
Best Local Similarity 62.5%; Score 28; DB 5; Length 198;
Best Local Similarity 62.5%; Pred. No. 45;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NMLSEVER 8
Db 108 NMLSELEK 115

RESULT 10
US-10-198-070-46
; Sequence 46, Application US/10198070
; GENERAL INFORMATION:
; APPLICANT: AVERBACK, PAUL
; APPLICANT: GEMMELL, JACK
; TITLE OF INVENTION: PEPTIDES EFFECTIVE IN THE TREATMENT OF TUMORS AND OTHER
; TITLE OF INVENTION: CONDITIONS REQUIRING THE REMOVAL OR DESTRUCTION OF
; FILE REFERENCE: 59003.000008
; CURRENT APPLICATION NUMBER: US/10/198,070
; PRIOR FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: 60/306,161
; PRIOR FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: 60/306,150
; PRIOR FILING DATE: 2001-07-19
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: 60/331,477
; NUMBER OF SEQ ID NOS: 125
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 46
; LENGTH: 244
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-198-070-46

Query Match
Best Local Similarity 100.0%; Score 28; DB 6; Length 244;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NMLSEV 6
Db 79 NMLSEV 84

RESULT 11
US-10-092-411A-3158
; Sequence 3158, Application US/10092411A
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032/96-101
; CURRENT APPLICATION NUMBER: US/10/092,411A
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 09/134,001
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5676
; SEQ ID NO 3158
; LENGTH: 344
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-10-092-411A-3158

Query Match
Best Local Similarity 55.6%; Score 28; DB 6; Length 344;
Best Local Similarity 55.6%; Pred. No. 83;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NMLSEVER 9
Db 300 NMLSELEKE 308

RESULT 12
PCT-US02-32032-16
; Sequence 16, Application PC/TUS0232032
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: YUE, Henry
; APPLICANT: WARREN, Bridget A.
; APPLICANT: LEHR-MASON, Patricia M.
; APPLICANT: TRAN, Uyen K.
; APPLICANT: DUGGAN, Brendan M.
; APPLICANT: THANGAVELU, Kavitha
; APPLICANT: YANG, Junming
; APPLICANT: XU, Yuming
; APPLICANT: TANG, Y. Tom
; APPLICANT: CHAMLA, Narinder K.
; APPLICANT: ELIOTT, Vicki S.
; APPLICANT: FORSYTHE, Ian J.
; APPLICANT: BECHA, Shanya D.
; APPLICANT: YAO, Monique G.
; APPLICANT: EMERLING, Brooke M.
; APPLICANT: GRIFFIN, Jennifer A.
; APPLICANT: LAL, Preeti G.
; APPLICANT: ZEBARJADIAN, Yeganeh
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: LEE, Ernestine A.
; APPLICANT: LEE, Soo Yeun
; APPLICANT: ELIOTT, Vicki S.
; APPLICANT: RAKKUMAR, Jayalaxmi
; APPLICANT: GORVAD, Ann E.
; APPLICANT: KABLE, Amy E.
; APPLICANT: LU, Dyung Aina M.

```
APPLICANT: BOROWSKY, Mark L.
TITLE OF INVENTION: SECRETED PROTEINS
FILE REFERENCE: PF-1217 PCT
CURRENT APPLICATION NUMBER: PCT/US02/32032
CURRENT FILING DATE: 2002-10-03
PRIOR APPLICATION NUMBER: US 60/326,945
PRIOR FILING DATE: 2001-10-03
PRIOR APPLICATION NUMBER: US 60/343,718
PRIOR FILING DATE: 2001-10-19
PRIOR APPLICATION NUMBER: US 60/343,980
PRIOR FILING DATE: 2001-11-02
PRIOR APPLICATION NUMBER: US 60/332,426
PRIOR FILING DATE: 2001-11-16
NUMBER OF SEQ ID NOS: 64
SOFTWARE: PERL Program
SEQ ID NO 16
LENGTH: 429
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc-feature
OTHER INFORMATION: Incyte ID No: 7500178CD1
PCT-US02-32032-16
```

```
Query Match      65.1%  Score 28;  DB 1;  Length 429;
Best Local Similarity 66.7%  Pred. No. 1.1e+02;
Matches 6;  Conservative 1;  Mismatches 2;  Indels 0;  Gaps 0;
```

```
OY      1  NMLSEVERE 9
        1:111111
Db      158  NMLFEKERE 166
```

```
RESULT 13
US-10-131-813A-94
Sequence 94, Application US/10131813A
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: Deforge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C139
CURRENT APPLICATION NUMBER: US/10/131,813A
CURRENT FILING DATE: 2002-04-24
PRIOR APPLICATION NUMBER: 60/049911
PRIOR FILING DATE: 1997-06-18
PRIOR APPLICATION NUMBER: 60/056974
PRIOR FILING DATE: 1997-08-26
PRIOR APPLICATION NUMBER: 60/059113
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059115
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059117
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059122
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059184
PRIOR FILING DATE: 1997-09-17
```

```
PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059352
PRIOR FILING DATE: 1997-09-19
PRIOR APPLICATION NUMBER: 60/059588
PRIOR FILING DATE: 1997-09-19
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 94
LENGTH: 429
TYPE: PRT
ORGANISM: Homo Sapien
US-10-131-813A-94
```

```
Query Match      65.1%  Score 28;  DB 6;  Length 429;
Best Local Similarity 66.7%  Pred. No. 1.1e+02;
Matches 6;  Conservative 1;  Mismatches 2;  Indels 0;  Gaps 0;
```

```
OY      1  NMLSEVERE 9
        1:111111
Db      158  NMLFEKERE 166
```

```
RESULT 14
US-10-131-819A-94
Sequence 94, Application US/10131819A
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: Deforge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C134
CURRENT APPLICATION NUMBER: US/10/131,819A
CURRENT FILING DATE: 2002-04-24
PRIOR APPLICATION NUMBER: 60/049911
PRIOR FILING DATE: 1997-06-18
PRIOR APPLICATION NUMBER: 60/056974
PRIOR FILING DATE: 1997-08-26
PRIOR APPLICATION NUMBER: 60/059113
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059115
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059117
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059122
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059184
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059352
PRIOR FILING DATE: 1997-09-19
PRIOR APPLICATION NUMBER: 60/059588
PRIOR FILING DATE: 1997-09-19
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 94
LENGTH: 429
```

; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-131-819A-94

Query Match 65.1%; Score 28; DB 6; Length 429;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 NMLSEVERE 9
|:| | | | |
Db 158 NMLFEKERE 166

RESULT 15
US-10-131-823A-94
; Sequence 94, Application US/10131823A
; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Mei-Qiang
; APPLICANT: Gerltsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Collin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C143
; CURRENT APPLICATION NUMBER: US/10/131,823A
; PRIOR FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-19
; Remaining Prior Application data removed - See file wrapper or PALM.
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 94
; LENGTH: 429
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-131-823A-94

Db 158 NMLFEKERE 166
Search completed: November 18, 2002, 16:05:20
Job time : 3.19355 secs

Query Match 65.1%; Score 28; DB 6; Length 429;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
OY 1 NMLSEVERE 9
|:| | | | |

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using SW model

Run on: November 18, 2002, 15:53:56 ; Search time 145.129 Seconds
(without alignments)
48.867 Million cell updates/sec

Title: US-09-848-781-2
Perfect score: 64
Sequence: 1 ACCDEYLQTKR 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 64473110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Pending_Patents_AA_Main:*

1: /cgn2_6/ptodata/1/paa/PCRT05_COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US06_COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US07_COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US080_COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US082_COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US083_COMB.pep.*
8: /cgn2_6/ptodata/1/paa/US084_COMB.pep.*
9: /cgn2_6/ptodata/1/paa/US085_COMB.pep.*
10: /cgn2_6/ptodata/1/paa/US086_COMB.pep.*
11: /cgn2_6/ptodata/1/paa/US087_COMB.pep.*
12: /cgn2_6/ptodata/1/paa/US088_COMB.pep.*
13: /cgn2_6/ptodata/1/paa/US089_COMB.pep.*
14: /cgn2_6/ptodata/1/paa/US090_COMB.pep.*
15: /cgn2_6/ptodata/1/paa/US091_COMB.pep.*
16: /cgn2_6/ptodata/1/paa/US092_COMB.pep.*
17: /cgn2_6/ptodata/1/paa/US093_COMB.pep.*
18: /cgn2_6/ptodata/1/paa/US094_COMB.pep.*
19: /cgn2_6/ptodata/1/paa/US095_COMB.pep.*
20: /cgn2_6/ptodata/1/paa/US096_COMB.pep.*
21: /cgn2_6/ptodata/1/paa/US097_COMB.pep.*
22: /cgn2_6/ptodata/1/paa/US098_COMB.pep.*
23: /cgn2_6/ptodata/1/paa/US099_COMB.pep.*
24: /cgn2_6/ptodata/1/paa/US100_COMB.pep.*
25: /cgn2_6/ptodata/1/paa/US101_COMB.pep.*
26: /cgn2_6/ptodata/1/paa/US102_COMB.pep.*
27: /cgn2_6/ptodata/1/paa/US60_COMB.pep.*

Pred. NO. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	64	100.0	11	22	US-09-848-781-2
2	54	84.4	289	21	US-09-791-537-21294
3	54	84.4	289	21	US-09-791-537-37210
4	54	84.4	289	21	US-09-791-537-42310
5	45	70.3	1111	27	US-60-360-039-5280
6	41	64.1	77	24	US-10-029-386-29564

7	41	64.1	539	21	US-09-708-427-29416	Sequence 29416, A
8	41	64.1	549	21	US-09-708-427-29415	Sequence 29415, A
9	41	64.1	560	21	US-09-708-427-29414	Sequence 29414, A
10	40	62.5	18	22	US-09-864-761-47551	Sequence 47551, A
11	40	62.5	18	25	US-10-182-098-11790	Sequence 11790, A
12	40	62.5	18	26	US-10-203-135-26759	Sequence 26759, A
13	40	62.5	18	26	US-10-203-138-12101	Sequence 12101, A
14	40	62.5	18	26	US-10-203-139-26928	Sequence 26928, A
15	40	62.5	713	27	US-60-236-359-16562	Sequence 16562, A
16	40	62.5	713	27	US-09-708-427-27132	Sequence 27132, A
17	40	62.5	721	19	US-09-570-581A-997	Sequence 997, App
18	40	62.5	722	21	US-09-708-427-27131	Sequence 27131, A
19	40	62.5	767	21	US-09-708-427-27130	Sequence 27130, A
20	39	60.9	70	27	US-60-162-247-4041	Sequence 4041, Ap
21	39	60.9	70	27	US-60-163-123-1761	Sequence 1761, Ap
22	39	60.9	71	27	US-60-162-243-672	Sequence 672, App
23	39	60.9	71	27	US-60-162-247-5037	Sequence 5037, App
24	39	60.9	1425	27	US-60-212-664-563	Sequence 563, App
25	39	60.9	1438	1	PCT-US01-08631-33990	Sequence 33990, A
26	39	60.9	1438	1	PCT-US01-08631-58667	Sequence 58667, A
27	39	60.9	1439	1	PCT-US01-08631-45208	Sequence 45208, A
28	38	59.4	87	21	US-09-708-427-42506	Sequence 42506, A
29	38	59.4	219	20	US-09-614-150-41031	Sequence 41031, A
30	38	59.4	219	27	US-60-191-637-4006	Sequence 4066, A
31	38	59.4	219	27	US-60-191-700-600	Sequence 600, App
32	38	59.4	252	21	US-09-791-537-83850	Sequence 83850, A
33	38	59.4	367	21	US-09-791-537-102975	Sequence 102975, A
34	37	57.8	48	21	US-09-733-089-23874	Sequence 23874, A
35	37	57.8	48	22	US-09-816-660-23874	Sequence 23874, A
36	37	57.8	70	1	PCT-US01-08631-53771	Sequence 53771, A
37	37	57.8	89	19	US-09-513-996A-31209	Sequence 31209, A
38	37	57.8	89	19	US-09-513-996A-46453	Sequence 46453, A
39	37	57.8	89	20	US-09-620-393B-6593	Sequence 6593, App
40	37	57.8	92	19	US-09-513-996A-31208	Sequence 31208, A
41	37	57.8	92	19	US-09-513-996A-46452	Sequence 46452, A
42	37	57.8	92	20	US-60-620-393B-6592	Sequence 6592, App
43	37	57.8	95	19	US-09-513-996A-31207	Sequence 31207, A
44	37	57.8	95	19	US-09-513-996A-46451	Sequence 46451, A
45	37	57.8	95	20	US-09-620-393B-6591	Sequence 6591, App

ALIGNMENTS

RESULT 1

US-09-848-781-2

; Sequence 2, Application US/09848781

; GENERAL INFORMATION:

; APPLICANT: SANDERS, MITCHELL

; TITLE OF INVENTION: A DEVICE FOR DETECTING BACTERIAL CONTAMINATION AND METHOD OF U

; FILE REFERENCE: 102951-10

; CURRENT APPLICATION NUMBER: US/09/848,781

; PRIORITY FILING DATE: 2001-05-03

; PRIORITY APPLICATION NUMBER: 60/201,405

; PRIORITY FILING DATE: 2000-05-03

; NUMBER OF SEQ ID NOS: 5

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 2

; LENGTH: 11

; TYPE: PRT

; ORGANISM: Listeria monocytogenes

US-09-848-781-2

Query Match 100.0%; Score 64; DB 22; Length 11;
Best Local Similarity 100.0%; Pred. NO. 0.002; 0; Indels 0; Gaps 0;
Matches 11; Conservative 0; Mismatches 0;

QY 1 ACCDEYLQTKR 11
|||||
DB 1 ACCDEYLQTKR 11

RESULT 2

```
US-09-791-537-21294
; Sequence 21294, Application US/09/91537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 21294
; TYPE: PRT
; ORGANISM: Listeria monocytogenes
US-09-791-537-21294
```

```
Query Match
Best Local Similarity 84.4%; Score 54; DB 21; Length 289;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 ACCDEYLTQ 9
Db 27 ACCDEYLTQ 35
```

```
RESULT 3
US-09-791-537-37210
; Sequence 37210, Application US/09/91537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 37210
; LENGTH: 289
; TYPE: PRT
; ORGANISM: Listeria monocytogenes
US-09-791-537-37210
```

```
Query Match
Best Local Similarity 84.4%; Score 54; DB 21; Length 289;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 ACCDEYLTQ 9
Db 27 ACCDEYLTQ 35
```

```
RESULT 4
US-09-791-537-42310
; Sequence 42310, Application US/09/91537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 42310
; LENGTH: 289
```

```
; TYPE: PRT
; ORGANISM: Listeria monocytogenes
US-09-791-537-42310
```

```
Query Match
Best Local Similarity 84.4%; Score 54; DB 21; Length 289;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 ACCDEYLTQ 9
Db 27 ACCDEYLTQ 35
```

```
RESULT 5
US-60-360-039-5280
; Sequence 5280, Application US/60360039
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Chen, Xianfeng
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)A
; CURRENT APPLICATION NUMBER: US/60/360,039
; CURRENT FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 5280
; LENGTH: 1111
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-60-360-039-5280
```

```
Query Match
Best Local Similarity 70.3%; Score 45; DB 27; Length 1111;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
```

```
OY 2 CCDEYLTQKE 11
Db 105 CCDEYLTQKED 114
```

```
RESULT 6
US-10-029-386-29564
; Sequence 29564, Application US/10029386
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: A60MICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 29564
; LENGTH: 77
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR17.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.63
; OTHER INFORMATION: SWISSPROT HIT: P02446, EVALU6 5.00e-16
US-10-029-386-29564
```

```
Query Match
Best Local Similarity 64.1%; Score 41; DB 24; Length 77;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
OY 1 ACCDEYLTQ 8
; 111 111
```


PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annotmax Sequence Listing Engine vers. 1.1
SEQ ID NO 47351
LENGTH: 18
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC011331.1
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.8
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.6
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.2
US-09-864-761-47551

Query Match 62.5%; Score 40; DB 22; Length 18;
Best Local Similarity 63.6%; Pred. No. 29;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 ACCDEYLOTKE 11
||||: |||
DB 4 ACCDDLTLTAE 14

RESULT 11
US-10-182-998-11790
Sequence 11790, Application US/10182998
GENERAL INFORMATION:
APPLICANT: Molecular Dynamics, Inc.
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
TITLE OF INVENTION: ANALYSIS OF GENE EXPRESSION IN HBL 100
FILE REFERENCE: PB 0004 WO 9
CURRENT APPLICATION NUMBER: US/10/182,998
PRIOR FILING DATE: 2002-08-02
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 04 February 2000 (04.02.00)
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 26 May 2000 (26.05.00)
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 03 August 2000 (03.08.00)
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 03 October 2000 (03.10.00)
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 27 September 2000 (27.09.00)
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 21 September 2000 (21.09.00)
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 30 June 2000 (30.06.00)
NUMBER OF SEQ ID NOS: 15009
SOFTWARE: Molecular Dynamics Sequence Listing Engine
SEQ ID NO 11790
LENGTH: 18
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC011331.1
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.2
US-10-182-998-11790

Query Match 62.5%; Score 40; DB 25; Length 18;
Best Local Similarity 63.6%; Pred. No. 29;

Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 1 ACCDEYLOTKE 11
||||: |||
DB 4 ACCDDLTLTAE 14

RESULT 12
US-10-203-135-26759
Sequence 26759, Application US/10203135
GENERAL INFORMATION:
APPLICANT: Molecular Dynamics, Inc.
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
TITLE OF INVENTION: ANALYSIS OF GENE EXPRESSION IN HUMAN LUNG
FILE REFERENCE: PB 0004 WO 5
CURRENT APPLICATION NUMBER: US/10/203,135
PRIOR FILING DATE: 2002-08-02
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 04 February 2000 (04.02.00)
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 26 May 2000 (26.05.00)
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 03 August 2000 (03.08.00)
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 03 October 2000 (03.10.00)
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 27 September 2000 (27.09.00)
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 21 September 2000 (21.09.00)
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 30 June 2000 (30.06.00)
NUMBER OF SEQ ID NOS: 37012
SOFTWARE: Molecular Dynamics Sequence Listing Engine
SEQ ID NO 26759
LENGTH: 18
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC011331.1
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.1
US-10-203-135-26759

Query Match 62.5%; Score 40; DB 26; Length 18;
Best Local Similarity 63.6%; Pred. No. 29;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 ACCDEYLOTKE 11
||||: |||
DB 4 ACCDDLTLTAE 14

RESULT 13
US-10-203-138-12101
Sequence 12101, Application US/10203138
GENERAL INFORMATION:
APPLICANT: Molecular Dynamics, Inc.
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
TITLE OF INVENTION: ANALYSIS OF GENE EXPRESSION IN BT 474
FILE REFERENCE: PB 0004 WO 8
CURRENT APPLICATION NUMBER: US/10/203,138
PRIOR FILING DATE: 2002-08-02
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 04 February 2000 (04.02.00)
PRIOR APPLICATION NUMBER: US 60/207,456

; PRIOR FILING DATE: 26 May 2000 (26.05.00)
 ; PRIOR APPLICATION NUMBER: US 09/532,366
 ; PRIOR FILING DATE: 03 August 2000 (03.08.00)
 ; PRIOR APPLICATION NUMBER: GB 24263.6
 ; PRIOR FILING DATE: 03 October 2000 (03.10.00)
 ; PRIOR APPLICATION NUMBER: US 60/236,359
 ; PRIOR FILING DATE: 27 September 2000 (27.09.00)
 ; PRIOR APPLICATION NUMBER: US 60/234,687
 ; PRIOR FILING DATE: 21 September 2000 (21.09.00)
 ; PRIOR APPLICATION NUMBER: US 09/608,408
 ; PRIOR FILING DATE: 30 June 2000 (30.06.00)
 ; NUMBER OF SEQ ID NOS: 15438
 ; SOFTWARE: Molecular Dynamics Sequence Listing Engine
 ; SEQ ID NO 12101
 ; LENGTH: 18
 ; TYPE: prt
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: MAP TO AC011331.1
 ; US-10-203-138-12101
 ; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.8

Query Match
 Best Local Similarity 62.5%; Score 40; DB 26; Length 18;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 ACCDEYLQTK 11
 DB 4 ACCDDLTLTAE 14

RESULT 14
 US-10-203-139-26928
 ; Sequence 26928, Application US/10203139
 ; GENERAL INFORMATION:
 ; APPLICANT: Molecular Dynamics, Inc.
 ; APPLICANT: Penn, Sharon G.
 ; APPLICANT: Rank, David R.
 ; APPLICANT: Hanzel, David K.
 ; APPLICANT: Chen, Wensheng
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
 ; FILE REFERENCE: PB 0004 WO 4
 ; CURRENT APPLICATION NUMBER: US/10/203,139
 ; PRIOR FILING DATE: 2002-08-02
 ; PRIOR APPLICATION NUMBER: US 60/180,312
 ; PRIOR FILING DATE: 04 February 2000 (04.02.00)
 ; PRIOR APPLICATION NUMBER: US 60/207,456
 ; PRIOR FILING DATE: 26 May 2000 (26.05.00)
 ; PRIOR APPLICATION NUMBER: US 09/532,366
 ; PRIOR FILING DATE: 03 August 2000 (03.08.00)
 ; PRIOR APPLICATION NUMBER: GB 24263.6
 ; PRIOR FILING DATE: 03 October 2000 (03.10.00)
 ; PRIOR APPLICATION NUMBER: US 60/236,359
 ; PRIOR FILING DATE: 27 September 2000 (27.09.00)
 ; PRIOR APPLICATION NUMBER: US 60/234,687
 ; PRIOR FILING DATE: 21 September 2000 (21.09.00)
 ; PRIOR APPLICATION NUMBER: US 09/608,408
 ; PRIOR FILING DATE: 30 June 2000 (30.06.00)
 ; NUMBER OF SEQ ID NOS: 37156
 ; SOFTWARE: Molecular Dynamics Sequence Listing Engine
 ; SEQ ID NO 26928
 ; LENGTH: 18
 ; TYPE: prt
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: MAP TO AC011331.1
 ; US-10-203-139-26928
 ; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.6

Query Match 62.5%; Score 40; DB 26; Length 18;

Best Local Similarity 63.6%; Pred. No. 29;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 ACCDEYLQTK 11
 DB 4 ACCDDLTLTAE 14

RESULT 15
 US-60-236-359-16562
 ; Sequence 16562, Application US/60236359
 ; GENERAL INFORMATION:
 ; APPLICANT: Penn, Sharon G.
 ; APPLICANT: Rank, David R.
 ; APPLICANT: Hanzel, David K.
 ; APPLICANT: Chen, Wensheng
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL F
 ; FILE REFERENCE: MDHMOF-4P
 ; CURRENT APPLICATION NUMBER: US/60/236,359
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: US 60/207,456
 ; PRIOR FILING DATE: 2000-05-26
 ; NUMBER OF SEQ ID NOS: 21709
 ; SOFTWARE: Molecular Dynamics Sequence Listing Engine
 ; SEQ ID NO 16562
 ; LENGTH: 18
 ; TYPE: prt
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: MAP TO AC011331.1
 ; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.1
 ; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.8
 ; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.6
 ; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.2
 ; US-60-236-359-16562

Query Match 62.5%; Score 40; DB 27; Length 18;
 Best Local Similarity 63.6%; Pred. No. 29;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 ACCDEYLQTK 11
 DB 4 ACCDDLTLTAE 14

Search completed: November 18, 2002, 16:05:06
 Job time : 147.129 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 18, 2002, 15:55:36 ; Search time 3.90323 Seconds
(without alignments)
36.183 Million cell updates/sec

Title: US-09-848-781-2
Perfect score: 64
Sequence: 1 ACDEYLOTKE 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 47062 seqs, 12839170 residues

Total number of hits satisfying chosen parameters: 47062

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_New:*
1: /cgn2_6/ptodata/2/paa/PCF_NEW_COMB.pep:*
2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep:*
7: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	35	54.7	186	US-10-264-237-2581	Sequence 2581, Ap
2	34	53.1	117	US-09-513-999C-7841	Sequence 7841, Ap
3	33.5	52.3	87	PCF-US02-27671A-11	Sequence 11, Appl
4	33	51.6	446	US-10-092-411A-5125	Sequence 5125, Ap
5	32	50.0	53	US-09-513-999C-7763	Sequence 7763, Ap
6	31	48.4	1154	US-10-092-411A-3428	Sequence 3428, Ap
7	30	48.9	21	US-10-062-109A-741	Sequence 741, App
8	30	46.9	58	PCF-US02-31861-3	Sequence 3, Appl
9	30	46.9	58	US-10-264-480-3	Sequence 4, Appl
10	30	46.9	58	US-10-264-480-4	Sequence 4, Appl
11	30	46.9	67	US-10-001-221A-7	Sequence 7, Appl
12	30	46.9	68	US-10-001-221A-3	Sequence 3, Appl
13	30	46.9	116	US-10-131-813A-534	Sequence 534, App
14	30	46.9	116	US-10-131-819A-534	Sequence 534, App
15	30	46.9	116	US-10-131-823A-534	Sequence 534, App
16	30	46.9	116	US-10-131-823A-534	Sequence 534, App
17	30	46.9	116	US-10-131-824A-534	Sequence 534, App
18	30	46.9	116	US-10-131-825A-534	Sequence 534, App
19	30	46.9	116	US-10-131-829A-534	Sequence 534, App
20	30	46.9	116	US-10-125-926A-534	Sequence 534, App
21	30	46.9	116	US-10-127-829A-534	Sequence 534, App
22	30	46.9	116	US-10-127-831A-534	Sequence 534, App
23	30	46.9	116	US-10-127-835A-534	Sequence 534, App
24	30	46.9	116	US-10-127-837A-534	Sequence 534, App
25	30	46.9	116	US-10-127-842A-534	Sequence 534, App
26	30	46.9	116	US-10-127-850A-534	Sequence 534, App

27	30	46.9	116	US-10-127-901A-534	Sequence 534, App
28	30	46.9	116	US-10-128-689A-534	Sequence 534, App
29	30	46.9	116	US-10-131-830A-534	Sequence 534, App
30	30	46.9	116	US-10-131-833A-534	Sequence 534, App
31	30	46.9	116	US-10-131-837A-534	Sequence 534, App
32	30	46.9	116	US-10-125-930A-534	Sequence 534, App
33	30	46.9	116	US-10-127-825A-534	Sequence 534, App
34	30	46.9	116	US-10-127-838A-534	Sequence 534, App
35	30	46.9	116	US-10-127-843A-534	Sequence 534, App
36	30	46.9	116	US-10-127-849A-534	Sequence 534, App
37	30	46.9	116	US-10-128-684A-534	Sequence 534, App
38	30	46.9	116	US-10-128-685A-534	Sequence 534, App
39	30	46.9	116	US-10-128-686A-534	Sequence 534, App
40	30	46.9	116	US-10-128-690A-534	Sequence 534, App
41	30	46.9	116	US-10-128-693A-534	Sequence 534, App
42	30	46.9	116	US-10-131-821A-534	Sequence 534, App
43	30	46.9	116	US-10-131-836A-534	Sequence 534, App
44	30	46.9	116	US-10-137-872A-534	Sequence 534, App
45	30	46.9	116	US-10-137-873A-534	Sequence 534, App

ALIGNMENTS

```
RESULT 1
US-10-264-237-2581
; Sequence 2581, Application US/10264237
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P4131P1
; CURRENT APPLICATION NUMBER: US/10/264,237
; PRIOR FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/16450
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/205,515
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 2876
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 2581
; LENGTH: 186
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-264-237-2581

Query Match      54.7%  Score 35; DB 6; Length 186;
Best Local Similarity 55.6%  Pred. No. 13;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      2 CCDEYLOTK 10
      ||: ||
DB      140 CCESFLTK 148

RESULT 2
US-09-513-999C-7841
; Sequence 7841, Application US/09513999C
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Ducterly, A.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: 59 US2,REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent. pm
; SEQ ID NO 7841
; LENGTH: 117
; TYPE: PRT
; ORGANISM: Homo sapiens
```

FEATURE:
NAME/KEY: SIGNAL
LOCATION: -27...1
OTHER INFORMATION: score 3.6
OTHER INFORMATION: seq ACLCSLVKTIDQ/FE
US-09-513-999C-7841

Query Match 53.1%; Score 34; DB 5; Length 117;
Best Local Similarity 75.0%; Pred. No. 12;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 CDEVLQTK 10
DB 36 CDAYLQMK 43

RESULT 3
PCT-US02-27671A-11
Sequence 11, Application PC/TUS0227671A
GENERAL INFORMATION:
APPLICANT: The General Hospital Corporation
TITLE OF INVENTION: Bone Morphogenic Protein 2-Induced Genes
TITLE OF INVENTION: and Polypeptides, and Their Use in Diagnostic and
TITLE OF INVENTION: Therapeutic Methods
FILE REFERENCE: 00786/409W02
CURRENT APPLICATION NUMBER: PCT/US02/27671A
CURRENT FILING DATE: 2002-08-29
PRIOR APPLICATION NUMBER: US 60/316,838
PRIOR FILING DATE: 2001-08-31
PRIOR APPLICATION NUMBER: US 60/317,287
PRIOR FILING DATE: 2001-09-05
NUMBER OF SEQ ID NOS: 29
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11
LENGTH: 87
TYPE: PRT
ORGANISM: Mus musculus
PCT-US02-27671A-11

Query Match 52.3%; Score 33.5; DB 1; Length 87;
Best Local Similarity 77.8%; Pred. No. 11;
Matches 7; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 2 CC-DEVLQTK 9
DB 4 CCNDEVLAT 12

RESULT 4
US-10-092-411A-5125
Sequence 5125, Application US/10092411A
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 032796-101
CURRENT APPLICATION NUMBER: US/10/092,411A
CURRENT FILING DATE: 2002-03-07
PRIOR APPLICATION NUMBER: US 09/134,001
PRIOR FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5676
SEQ ID NO 5125
LENGTH: 446
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-10-092-411A-5125

Query Match 51.6%; Score 33; DB 6; Length 446;
Best Local Similarity 50.0%; Pred. No. 68;

Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 2 CCDEVLQTK 11
DB 132 CTDREXTEE 141

RESULT 5
US-09-513-999C-7763
Sequence 7763, Application US/09513999C
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Duclet, A.
APPLICANT: Giordano, J.Y.
TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
FILE REFERENCE: 59 US2 REG
CURRENT APPLICATION NUMBER: US/09/513,999C
CURRENT FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/122,487
PRIOR FILING DATE: 1999-02-26
NUMBER OF SEQ ID NOS: 36681
SOFTWARE: Patent.pm
SEQ ID NO 7763
LENGTH: 53
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: UNSURE
LOCATION: 28
OTHER INFORMATION: Xaa-Cys or Gly
FEATURE:
NAME/KEY: UNSURE
LOCATION: 30
OTHER INFORMATION: Xaa- * or Glu
FEATURE:
NAME/KEY: UNSURE
LOCATION: 48
OTHER INFORMATION: Xaa-Cys or Gly or Arg or Ser
US-09-513-999C-7763

Query Match 50.0%; Score 32; DB 5; Length 53;
Best Local Similarity 83.3%; Pred. No. 13;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CCDEVL 7
DB 10 CCDDLTL 15

RESULT 6
US-10-092-411A-3428
Sequence 3428, Application US/10092411A
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 032796-101
CURRENT APPLICATION NUMBER: US/10/092,411A
CURRENT FILING DATE: 2002-03-07
PRIOR APPLICATION NUMBER: US 09/134,001
PRIOR FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5676
SEQ ID NO 3428
LENGTH: 1154
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-10-092-411A-3428

Query Match 48.4%; Score 31; DB 6; Length 1154;

Best Local Similarity 71.4%; Pred. No. 3.8e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 5 EYLQKE 11
:|:||||
Db 992 QYIQKE 998

RESULT 7
US-10-062-109A-741
; Sequence 741, Application US/10062109A
; GENERAL INFORMATION:
; APPLICANT: Agensys
; APPLICANT: Challita-Bid, Pia M.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Paris, Mary
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Jakobovits, Aya
; TITLE OF INVENTION: Nucleic Acid and Corresponding Protein
; TITLE OF INVENTION: Entitled 161p2f10b Useful in Treatment and Detection of
; FILE REFERENCE: 51158-20062.01
; CURRENT APPLICATION NUMBER: US/10/062,109A
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: US 10/005,480
; PRIOR FILING DATE: 2001-11-07
; NUMBER OF SEQ ID NOS: 765
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 741
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-062-109A-741

Query Match 46.9%; Score 30; DB 6; Length 21;
Best Local Similarity 45.5%; Pred. No. 11;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 ACCDEYLQKE 11
:|:||||
Db 2 SCSDCLQKDKD 12

RESULT 8
PCT-US02-31861-3
; Sequence 3, Application PC/TUS0231861
; GENERAL INFORMATION:
; APPLICANT: Hammock, Bruce D.
; APPLICANT: Inceoglu, Bora
; TITLE OF INVENTION: ISOLATED POLYPEPTIDES AND COMPOSITIONS
; TITLE OF INVENTION: FROM THE VENOM OF P. TRANSVAALICUS AND METHODS OF USE
; FILE REFERENCE: UCAL256WO
; CURRENT APPLICATION NUMBER: PCT/US02/31861
; CURRENT FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: 60/393,070
; PRIOR FILING DATE: 2002-06-28
; PRIOR APPLICATION NUMBER: 60/327,602
; PRIOR FILING DATE: 2001-10-04
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 58
; TYPE: PRT
; ORGANISM: Parabuthus transvaalicus
PCT-US02-31861-3

Query Match 46.9%; Score 30; DB 1; Length 58;
Best Local Similarity 55.6%; Pred. No. 31;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 3 CDEYLQKE 11
:|:||||
:|:||||

Db 48 CKEXLDKDKD 56

RESULT 9
PCT-US02-31861-4
; Sequence 4, Application PC/TUS0231861
; GENERAL INFORMATION:
; APPLICANT: Hammock, Bruce D.
; APPLICANT: Inceoglu, Bora
; TITLE OF INVENTION: ISOLATED POLYPEPTIDES AND COMPOSITIONS
; TITLE OF INVENTION: FROM THE VENOM OF P. TRANSVAALICUS AND METHODS OF USE
; FILE REFERENCE: UCAL256WO
; CURRENT APPLICATION NUMBER: PCT/US02/31861
; CURRENT FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: 60/393,070
; PRIOR FILING DATE: 2002-06-28
; PRIOR APPLICATION NUMBER: 60/327,602
; PRIOR FILING DATE: 2001-10-04
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 58
; TYPE: PRT
; ORGANISM: Parabuthus transvaalicus
PCT-US02-31861-4

Query Match 46.9%; Score 30; DB 1; Length 58;
Best Local Similarity 55.6%; Pred. No. 31;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 CDEYLQKE 11
:|:||||
Db 48 CKEXLDKDKD 56

RESULT 10
US-10-264-480-3
; Sequence 3, Application US/10264480
; GENERAL INFORMATION:
; APPLICANT: Hammock, Bruce D.
; APPLICANT: Inceoglu, Bora
; TITLE OF INVENTION: ISOLATED POLYPEPTIDES AND COMPOSITIONS
; TITLE OF INVENTION: FROM THE VENOM OF P. TRANSVAALICUS AND METHODS OF USE
; FILE REFERENCE: UCAL256
; CURRENT APPLICATION NUMBER: US/10/264,480
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: 60/393,070
; PRIOR FILING DATE: 2002-06-28
; PRIOR APPLICATION NUMBER: 60/327,602
; PRIOR FILING DATE: 2001-10-04
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 58
; TYPE: PRT
; ORGANISM: Parabuthus transvaalicus
US-10-264-480-3

Query Match 46.9%; Score 30; DB 6; Length 58;
Best Local Similarity 55.6%; Pred. No. 31;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 CDEYLQKE 11
:|:||||
Db 48 CKEXLDKDKD 56

RESULT 11
US-10-264-480-4
; Sequence 4, Application US/10264480
; GENERAL INFORMATION:
; APPLICANT: Hammock, Bruce D.
; APPLICANT: Inceoglu, Bora

```
; TITLE OF INVENTION: ISOLATED POLYPEPTIDES AND COMPOSITIONS
; TITLE OF INVENTION: FROM THE VENOM OF P. TRANSVAALICUS AND METHODS OF USE
; FILE REFERENCE: UCAL256
; CURRENT APPLICATION NUMBER: US/10/264,480
; PRIOR FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: 60/393,070
; PRIOR FILING DATE: 2002-06-28
; PRIOR APPLICATION NUMBER: 60/327,602
; PRIOR FILING DATE: 2001-10-04
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 58
; TYPE: PRT
; ORGANISM: Parabuthus transvaalicus
; US-10-264-480-4

Query Match          46.9%; Score 30; DB 6; Length 58;
Best Local Similarity 55.6%; Pred. No. 31;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY      3 CDEYLTQKE 11
Db      48 CKETLDKDK 56

RESULT 12
; US-10-001-221A-7
; GENERAL INFORMATION:
; APPLICANT: Schall, Thomas J. Talbot, Dale Berkowitz, Robert
; APPLICANT: Zheng, Wei Premack, Brett Howard, Maureen
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INDUCING AN IMMUNE RESPONSE
; FILE REFERENCE: 10709/14
; CURRENT APPLICATION NUMBER: US/10/001,221A
; CURRENT FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 09/834,814
; PRIOR FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 67
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Chimeric molecule
; US-10-001-221A-7

Query Match          46.9%; Score 30; DB 6; Length 67;
Best Local Similarity 42.9%; Pred. No. 35;
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY      2 CCDEYLTQ 8
Db      11 CCQDYIR 17

RESULT 13
; US-10-001-221A-3
; GENERAL INFORMATION:
; APPLICANT: Schall, Thomas J. Talbot, Dale Berkowitz, Robert
; APPLICANT: Zheng, Wei Premack, Brett Howard, Maureen
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INDUCING AN IMMUNE RESPONSE
; FILE REFERENCE: 10709/14
; CURRENT APPLICATION NUMBER: US/10/001,221A
; CURRENT FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 09/834,814
; PRIOR FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 68
```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-001-221A-3

Query Match          46.9%; Score 30; DB 6; Length 68;
Best Local Similarity 42.9%; Pred. No. 36;
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY      2 CCDEYLTQ 8
Db      12 CCQDYIR 18

RESULT 14
; US-10-131-813A-534
; Sequence 534, Application US/10131813A
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330RIC139
; CURRENT APPLICATION NUMBER: US/10/131,813A
; CURRENT FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-19
; Remaining prior Application data removed - See File Wrapper or PAM.
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 534
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Homo Saplen
; US-10-131-813A-534

Query Match          46.9%; Score 30; DB 6; Length 116;
Best Local Similarity 44.4%; Pred. No. 61;
Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY      2 CDEYLTQTK 10
Db      11 :::::1:
```

Db 69 CCYDHLKTQ 77

RESULT 15

US-10-131-819A-534

; Sequence 534, Application US/10131819A

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Beresini, Maureen

; APPLICANT: Deforge, Laura

; APPLICANT: Desnoyers, Luc

; APPLICANT: Filvaroff, Ellen

; APPLICANT: Gao, Wei-Qiang

; APPLICANT: Gerritsen, Mary E.

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Sherwood, Steven

; APPLICANT: Smith, Victoria

; APPLICANT: Stewart, Timothy A.

; APPLICANT: Tumas, Daniel

; APPLICANT: Watanabe, Colin K

; APPLICANT: Wood, William

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; FILE OF INVENTION: ACIDS ENCODING THE SAME

; FILE REFERENCE: P3330R1C134

; CURRENT APPLICATION NUMBER: US/10/131,819A

; CURRENT FILING DATE: 2002-04-24

; PRIOR APPLICATION NUMBER: 60/049911

; PRIOR FILING DATE: 1997-06-18

; PRIOR APPLICATION NUMBER: 60/056974

; PRIOR FILING DATE: 1997-08-26

; PRIOR APPLICATION NUMBER: 60/059113

; PRIOR FILING DATE: 1997-09-17

; PRIOR APPLICATION NUMBER: 60/059115

; PRIOR FILING DATE: 1997-09-17

; PRIOR APPLICATION NUMBER: 60/059117

; PRIOR FILING DATE: 1997-09-17

; PRIOR APPLICATION NUMBER: 60/059122

; PRIOR FILING DATE: 1997-09-17

; PRIOR APPLICATION NUMBER: 60/059184

; PRIOR FILING DATE: 1997-09-17

; PRIOR APPLICATION NUMBER: 60/059263

; PRIOR FILING DATE: 1997-09-18

; PRIOR APPLICATION NUMBER: 60/059352

; PRIOR FILING DATE: 1997-09-19

; PRIOR APPLICATION NUMBER: 60/059588

; PRIOR FILING DATE: 1997-09-19

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 550

; SEQ ID NO 534

; LENGTH: 116

; TYPE: PRT

; ORGANISM: Homo Sapien

US-10-131-819A-534

Query Match 46.9%; Score 30; DB 6; Length 116;

Best Local Similarity 44.4%; Pred. No. 61;

Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Oy 2 CCDEYLQTK 10

Db 69 CCYDHLKTQ 77

Search completed: November 18, 2002, 16:05:21
Job time : 4.90323 secs

THIS PAGE BLANK (USPTO)

Mon Nov 25 10:11:48 2002

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Comugen Ltd.

OM protein - protein search, using sw model

Run on: November 18, 2002, 15:53:56 ; Search time 145.129 Seconds
(without alignments)
48.867 Million cell updates/sec

Title: US-09-848-781-3
Perfect score: 56
Sequence: 1 ADVEPTGAKE 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Pending_Patents_AA_Main:*

1: /cgn2_6/ptodata/1/paa/PCTUS_COMB.pep:*
2: /cgn2_6/ptodata/1/paa/US06_COMB.pep:*
3: /cgn2_6/ptodata/1/paa/US07_COMB.pep:*
4: /cgn2_6/ptodata/1/paa/US08_COMB.pep:*
5: /cgn2_6/ptodata/1/paa/US081_COMB.pep:*
6: /cgn2_6/ptodata/1/paa/US082_COMB.pep:*
7: /cgn2_6/ptodata/1/paa/US083_COMB.pep:*
8: /cgn2_6/ptodata/1/paa/US084_COMB.pep:*
9: /cgn2_6/ptodata/1/paa/US085_COMB.pep:*
10: /cgn2_6/ptodata/1/paa/US086_COMB.pep:*
11: /cgn2_6/ptodata/1/paa/US087_COMB.pep:*
12: /cgn2_6/ptodata/1/paa/US088_COMB.pep:*
13: /cgn2_6/ptodata/1/paa/US089_COMB.pep:*
14: /cgn2_6/ptodata/1/paa/US090_COMB.pep:*
15: /cgn2_6/ptodata/1/paa/US091_COMB.pep:*
16: /cgn2_6/ptodata/1/paa/US092_COMB.pep:*
17: /cgn2_6/ptodata/1/paa/US093_COMB.pep:*
18: /cgn2_6/ptodata/1/paa/US094_COMB.pep:*
19: /cgn2_6/ptodata/1/paa/US095_COMB.pep:*
20: /cgn2_6/ptodata/1/paa/US096_COMB.pep:*
21: /cgn2_6/ptodata/1/paa/US097_COMB.pep:*
22: /cgn2_6/ptodata/1/paa/US098_COMB.pep:*
23: /cgn2_6/ptodata/1/paa/US099_COMB.pep:*
24: /cgn2_6/ptodata/1/paa/US100_COMB.pep:*
25: /cgn2_6/ptodata/1/paa/US101_COMB.pep:*
26: /cgn2_6/ptodata/1/paa/US102_COMB.pep:*
27: /cgn2_6/ptodata/1/paa/US60_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	56	100.0	11	22	US-09-848-781-3
2	46	82.1	479	20	US-09-626-813A-63
3	46	82.1	479	23	US-09-971-309-64
4	46	82.1	1603	20	US-09-626-813A-62
5	44	78.6	289	21	US-09-791-537-21294
6	44	78.6	289	21	US-09-791-537-37210

7	44	78.6	289	21	US-09-791-537-42310	Sequence 42310, A
8	479	71.4	479	21	US-09-791-537-27807	Sequence 27807, A
9	38	67.9	80	18	US-09-417-507-39265	Sequence 39265, A
10	38	67.9	395	21	US-09-738-626-4698	Sequence 4698, Ap
11	37	66.1	199	21	US-09-791-537-110144	Sequence 110144, A
12	37	66.1	236	21	US-09-791-537-1011	Sequence 1011, Ap
13	37	66.1	226	21	US-09-741-669-381	Sequence 381, App
14	37	66.1	226	21	US-09-791-537-102286	Sequence 102286, A
15	37	66.1	247	18	US-09-489-039A-7921	Sequence 7921, Ap
16	37	66.1	374	20	US-09-629-469A-18558	Sequence 18558, A
17	37	66.1	459	17	PCT-US02-03987-11954	Sequence 11954, A
18	37	66.1	459	22	US-09-815-242-11954	Sequence 11954, A
19	37	66.1	459	24	US-10-072-851-11954	Sequence 11954, A
20	37	66.1	471	16	US-09-252-991A-18083	Sequence 18083, A
21	37	66.1	735	27	US-09-360-039-8229	Sequence 8229, A
22	37	66.1	846	27	US-09-360-039-9102	Sequence 9102, Ap
23	37	66.1	2719	1	PCT-US01-08656-8996	Sequence 8996, Ap
24	37	66.1	2719	1	PCT-US01-14827-13244	Sequence 13244, A
25	36	64.3	60	1	PCT-US01-08631-40216	Sequence 40216, A
26	36	64.3	127	1	PCT-US01-01354-16715	Sequence 16715, A
27	36	64.3	127	21	US-09-764-905-16715	Sequence 16715, A
28	36	64.3	127	24	US-10-092-399-16715	Sequence 16715, A
29	36	64.3	127	26	US-10-221-279-10748	Sequence 10748, A
30	36	64.3	158	21	US-09-758-471-4407	Sequence 4407, Ap
31	36	64.3	158	26	US-10-235-953-4407	Sequence 4407, Ap
32	36	64.3	359	20	US-09-614-150-15702	Sequence 15702, A
33	36	64.3	369	27	US-09-191-637-15746	Sequence 15746, A
34	36	64.3	369	27	US-09-191-681-12479	Sequence 12479, A
35	36	64.3	390	15	US-09-107-537-6615	Sequence 6615, Ap
36	36	64.3	390	15	US-09-107-532A-6615	Sequence 6615, Ap
37	36	64.3	534	21	US-09-739-449-9641	Sequence 9641, Ap
38	36	64.3	560	27	US-09-803-110-9641	Sequence 9641, Ap
39	36	64.3	560	27	US-09-167-217-15748	Sequence 15748, A
40	36	64.3	560	27	US-09-173-464-12911	Sequence 12911, A
41	36	64.3	642	21	US-09-791-537-146122	Sequence 146122, A
42	36	64.3	913	27	US-09-791-537-13309	Sequence 13309, A
43	36	64.3	1925	26	US-10-205-032-12	Sequence 12, Appl
44	36	64.3	2432	1	PCT-US01-08633-36382	Sequence 36382, A
45	36	64.3	4834	24	US-10-097-534-27	Sequence 27, Appl

ALIGNMENTS

RESULT 1
US-09-848-781-3
Sequence 3, Application US/09848781
GENERAL INFORMATION:
APPLICANT: SANDERS, MITCHELL
TITLE OF INVENTION: A DEVICE FOR DETECTING BACTERIAL CONTAMINATION AND METHOD OF U
FILE REFERENCE: 102951-10
CURRENT APPLICATION NUMBER: US/09/848,781
PRIOR FILING DATE: 2001-05-03
PRIOR APPLICATION NUMBER: 60/201,405
PRIOR FILING DATE: 2000-05-03
NUMBER OF SEQ ID NOS: 5
SOFTWARE: PatentIn version 3.0
SEQ ID NO 3
LENGTH: 11
TYPE: PRT
ORGANISM: *Listeria monocytogenes*
US-09-848-781-3

Query Match 100.0%; Score 56; DB 22; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.0013;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ADVEPTGAKE 11
|||||
Db 1 ADVEPTGAKE 11

RESULT 2

US-09-626-813A-63
; Sequence 63, Application US/09626813A
; GENERAL INFORMATION:
; APPLICANT: STRATAGENE
; TITLE OF INVENTION: Pfu REPLICATION ACCESSORY FACTORS AND METHODS OF USE
; FILE REFERENCE: 04121.0161-00000
; CURRENT APPLICATION NUMBER: US/09/626,813A
; PRIOR FILING DATE: 2000-07-27
; PRIOR FILING DATE: 1999-07-30
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 63
; LENGTH: 479
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Recombinant
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (354)
; OTHER INFORMATION: Xaa is any amino acid
US-09-626-813A-63

Query Match

82.1%; Score 46; DB 20; Length 479;

Best Local Similarity 72.7%; Pred. No. 10;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;OY 1 ADIYEPGAK 11
11 :11:1111

Db 110 ADNIPEGAK 120

RESULT 3

US-09-971-309-64
; Sequence 64, Application US/09971309
; GENERAL INFORMATION:
; APPLICANT: UEMORI, Takashi
; APPLICANT: SATO, Yoshimi
; APPLICANT: FUJITA, Tomoko
; APPLICANT: MIYAKE, Kazuo
; APPLICANT: MUKAI, Hiroyuki
; APPLICANT: ASADA, Kiyozo
; APPLICANT: KATO, Ikunoshin
; TITLE OF INVENTION: DNA POLYMERASE-RELATED FACTORS
; FILE REFERENCE: 1422-049AP
; CURRENT APPLICATION NUMBER: US/09/971,309
; CURRENT FILING DATE: 2001-10-05
; PRIOR FILING DATE: 1999-12-23
; PRIOR FILING DATE: 1998-06-24
; PRIOR FILING DATE: 1997-06-26
; PRIOR FILING DATE: 1997-11-27
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 64
; LENGTH: 479
; TYPE: PRT
; ORGANISM: Pyrococcus furiosus
US-09-971-309-64

Query Match

82.1%; Score 46; DB 23; Length 479;

Best Local Similarity 72.7%; Pred. No. 10;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;OY 1 ADIYEPGAK 11
11 :11:1111

Db 110 ADNIPEGAK 120

RESULT 4

US-09-626-813A-62
; Sequence 62, Application US/09626813A
; GENERAL INFORMATION:
; APPLICANT: STRATAGENE
; TITLE OF INVENTION: Pfu REPLICATION ACCESSORY FACTORS AND METHODS OF USE
; FILE REFERENCE: 04121.0161-00000
; CURRENT APPLICATION NUMBER: US/09/626,813A
; PRIOR FILING DATE: 2000-07-27
; PRIOR FILING DATE: 1999-07-30
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 62
; LENGTH: 1603
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Amino acid
; FEATURE:
; OTHER INFORMATION: sequence of the genomic RFC clone
; NAME/KEY: MISC_FEATURE
; LOCATION: (77)
; OTHER INFORMATION: Xaa is any amino acid
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (114)
; OTHER INFORMATION: Xaa is any amino acid
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (35)
; OTHER INFORMATION: Xaa is any amino acid
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (45)
; OTHER INFORMATION: Xaa is any amino acid
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (1269)
; OTHER INFORMATION: Xaa is any amino acid
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (1496)
; OTHER INFORMATION: Xaa is any amino acid
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (1491)
; OTHER INFORMATION: Xaa is any amino acid
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (1504)
; OTHER INFORMATION: Xaa is any amino acid
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (1519)
; OTHER INFORMATION: Xaa is any amino acid
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (1522)
; OTHER INFORMATION: Xaa is any amino acid
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (1542)
; OTHER INFORMATION: Xaa is any amino acid
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (1559)
; OTHER INFORMATION: Xaa is any amino acid
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (1570)
; OTHER INFORMATION: Xaa is any amino acid

```
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (1574)
; OTHER INFORMATION: Xaa is any amino acid
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (1581)
; OTHER INFORMATION: Xaa is any amino acid
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (1586)
; OTHER INFORMATION: Xaa is any amino acid
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (1591)
; OTHER INFORMATION: Xaa is any amino acid
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (1593)
; OTHER INFORMATION: Xaa is any amino acid
US-09-626-813A-62
```

```
Query Match      82.1%; Score 46; DB 20; Length 1603;
Best Local Similarity 72.7%; Pred. No. 43;
Matches      8; Conservative      2; Mismatches      1; Indels      0; Gaps      0;
```

```
QY      1 ADTVEPTGAK 11
      11 : ||: ||||
Db      1025 ADNIETPGAK 1035
```

```
RESULT 5
US-09-791-537-21294
; Sequence 21294, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 21294
; LENGTH: 289
; TYPE: PRT
; ORGANISM: Listeria monocytogenes
US-09-791-537-21294
```

```
Query Match      78.6%; Score 44; DB 21; Length 289;
Best Local Similarity 88.9%; Pred. No. 13;
Matches      8; Conservative      1; Mismatches      0; Indels      0; Gaps      0;
```

```
QY      2 DIVEPTGAK 10
      11 : |||||
Db      260 DIVEPTGAR 268
```

```
RESULT 6
US-09-791-537-37210
; Sequence 37210, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
```

```
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 37210
; LENGTH: 289
; TYPE: PRT
; ORGANISM: Listeria monocytogenes
US-09-791-537-37210
```

```
Query Match      78.6%; Score 44; DB 21; Length 289;
Best Local Similarity 88.9%; Pred. No. 13;
Matches      8; Conservative      1; Mismatches      0; Indels      0; Gaps      0;
```

```
QY      2 DIVEPTGAK 10
      11 : |||||
Db      260 DIVEPTGAR 268
```

```
RESULT 7
US-09-791-537-42310
; Sequence 42310, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 42310
; LENGTH: 289
; TYPE: PRT
; ORGANISM: Listeria monocytogenes
US-09-791-537-42310
```

```
Query Match      78.6%; Score 44; DB 21; Length 289;
Best Local Similarity 88.9%; Pred. No. 13;
Matches      8; Conservative      1; Mismatches      0; Indels      0; Gaps      0;
```

```
QY      2 DIVEPTGAK 10
      11 : |||||
Db      260 DIVEPTGAR 268
```

```
RESULT 8
US-09-791-537-27807
; Sequence 27807, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 27807
; LENGTH: 479
; TYPE: PRT
; ORGANISM: Pyrococcus abyssi
US-09-791-537-27807
```

```
Query Match      71.4%; Score 40; DB 21; Length 479;
Best Local Similarity 63.6%; Pred. No. 1,4e+02;
Matches      7; Conservative      2; Mismatches      2; Indels      0; Gaps      0;
```

```
QY      1 ADTVEPTGAK 11
      11 : ||: |||
Db      110 ADNIETPGAK 120
```

```
RESULT 9
; US-09-417-507-39265
; Sequence 39265, Application US/09417507
; GENERAL INFORMATION:
; APPLICANT: KEITH G. WEINSTOCK ET AL.
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCES RELATING TO ASPERGILLUS
; FILE REFERENCE: PAT99-10
; CURRENT APPLICATION NUMBER: US/09/417,507
; CURRENT FILING DATE: 1999-10-14
; NUMBER OF SEQ ID NOS: 44312
; LENGTH: 80
; TYPE: PRT
; ORGANISM: A.fumigatus
US-09-417-507-39265

Query Match      67.9%; Score 38; DB 18; Length 80;
Best Local Similarity 80.0%; Pred. No. 41;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ADVEPTGAK 10
   || || || ||
Db 32 ADVEATGAK 41

RESULT 10
; US-09-738-626-4698
; Sequence 4698, Application US/09738626
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOBUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHINI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 4698
; LENGTH: 395
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-4698

Query Match      67.9%; Score 38; DB 21; Length 395;
Best Local Similarity 80.0%; Pred. No. 2,8e+02;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ADVEPTGAK 10
   || || || || ||
Db 288 ADVEPTGAK 297

RESULT 11
; US-09-791-537-110144
; Sequence 110144, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.

; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 110144
; LENGTH: 199
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-791-537-110144

Query Match      66.1%; Score 37; DB 21; Length 199;
Best Local Similarity 80.0%; Pred. No. 1,9e+02;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ADVEPTGAK 10
   || || || ||
Db 69 ADVEATGAK 78

RESULT 12
; US-09-791-537-1011
; Sequence 1011, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1011
; LENGTH: 210
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-791-537-1011

Query Match      66.1%; Score 37; DB 21; Length 210;
Best Local Similarity 80.0%; Pred. No. 2e+02;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ADVEPTGAK 10
   || || || ||
Db 43 ADVEATGAK 52

RESULT 13
; US-09-741-669-381
; Sequence 381, Application US/09741669
; GENERAL INFORMATION:
; APPLICANT: Forsyth, R. Allyn
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; TITLE OF INVENTION: Genes identified as required for
; FILE REFERENCE: ELITRA.009A
; CURRENT APPLICATION NUMBER: US/09/741,669
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 60/173005
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 381
; LENGTH: 236
; TYPE: PRT
; ORGANISM: Escherichia coli
```


US-09-741-669-381

Query Match 66.1%; Score 37; DB 21; Length 236;
Best Local Similarity 80.0%; Pred. No. 2.4e+02;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ADTVEPTGAK 10
|| || ||||
DB 69 ADLVEATGAK 78

RESULT 14

US-09-791-537-102286
; Sequence 102286, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: Patent version 3.0
; SEQ ID NO 102286
; LENGTH: 236
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-791-537-102286

Query Match 66.1%; Score 37; DB 21; Length 236;
Best Local Similarity 80.0%; Pred. No. 2.4e+02;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ADTVEPTGAK 10
|| || ||||
DB 69 ADLVEATGAK 78

RESULT 15

US-09-489-039A-7921
; Sequence 7921, Application US/09489039A
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709,2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 7921
; LENGTH: 247
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-7921

Query Match 66.1%; Score 37; DB 18; Length 247;
Best Local Similarity 80.0%; Pred. No. 2.5e+02;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ADTVEPTGAK 10
|| || ||||
DB 79 ADLVEATGAK 88

Search completed: November 18, 2002, 16:05:07
Job time : 146.129 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 18, 2002, 15:55:36 ; Search time 3.90323 Seconds
(without alignments)
36.183 Million cell updates/sec

Title: US-09-848-781-3
Perfect score: 56
Sequence: 1 ADTYPEPTGAKK 11

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 47062 seqs, 12839170 residues

Total number of hits satisfying chosen parameters: 47062

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA.New:*
1: /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pep:*
2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep:*
7: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	31	55.4	270	1	PCT-US02-32852-4
2	31	53.4	424	6	US-10-092-411A-3876
3	31	55.4	435	6	US-10-266-252-14
4	31	55.4	529	6	US-10-125-923A-472
5	29	51.8	14	5	US-09-555-534B-17
6	29	51.8	86	5	US-09-555-534B-2
7	29	51.8	86	5	US-09-555-534B-4
8	29	51.8	86	5	US-09-555-534B-6
9	29	51.8	86	5	US-09-555-534B-6
10	29	51.8	148	6	US-10-092-411A-3835
11	29	51.8	180	5	US-09-905-666A-76
12	29	51.8	212	5	US-09-905-666A-64
13	29	51.8	212	5	US-09-905-666A-71
14	29	51.8	212	5	US-09-905-666A-72
15	29	51.8	213	5	US-09-905-666A-74
16	29	51.8	215	5	US-09-905-666A-67
17	29	51.8	215	5	US-09-905-666A-73
18	29	51.8	432	6	US-10-085-198-78
19	29	51.8	490	5	US-09-906-777B-96
20	29	51.8	490	5	US-09-904-011C-96
21	29	51.8	490	6	US-10-131-813A-356
22	29	51.8	490	6	US-10-131-819A-356
23	29	51.8	490	6	US-10-131-823A-356
24	29	51.8	490	6	US-10-131-824A-356
25	29	51.8	490	6	US-10-131-826A-356
26	29	51.8	490	6	US-10-131-829A-356

27	29	51.8	490	6	US-10-125-926A-356	Sequence 356, App
28	29	51.8	490	6	US-10-127-829A-356	Sequence 356, App
29	29	51.8	490	6	US-10-127-831A-356	Sequence 356, App
30	29	51.8	490	6	US-10-127-835A-356	Sequence 356, App
31	29	51.8	490	6	US-10-127-837A-356	Sequence 356, App
32	29	51.8	490	6	US-10-127-842A-356	Sequence 356, App
33	29	51.8	490	6	US-10-127-850A-356	Sequence 356, App
34	29	51.8	490	6	US-10-127-901A-356	Sequence 356, App
35	29	51.8	490	6	US-10-128-689A-356	Sequence 356, App
36	29	51.8	490	6	US-10-131-830A-356	Sequence 356, App
37	29	51.8	490	6	US-10-131-833A-356	Sequence 356, App
38	29	51.8	490	6	US-10-131-837A-356	Sequence 356, App
39	29	51.8	490	6	US-10-125-930A-356	Sequence 356, App
40	29	51.8	490	6	US-10-127-825A-356	Sequence 356, App
41	29	51.8	490	6	US-10-127-838B-356	Sequence 356, App
42	29	51.8	490	6	US-10-127-843A-356	Sequence 356, App
43	29	51.8	490	6	US-10-127-849A-356	Sequence 356, App
44	29	51.8	490	6	US-10-128-684A-356	Sequence 356, App
45	29	51.8	490	6	US-10-128-685A-356	Sequence 356, App

ALIGNMENTS

RESULT 1
PCT-US02-32852-4
Sequence 4, Application PC/TUS0232852
GENERAL INFORMATION:
APPLICANT: INCYTE GENOMICS, INC.
APPLICANT: TANG, Y. TOM
APPLICANT: FORSYTHE, Ian J.
APPLICANT: EMERLING, Brooke M.
APPLICANT: HAFALIA, April J.A.
APPLICANT: YUE, Henry
APPLICANT: XU, Yuning
APPLICANT: CIETZEN, Kimberly J.
APPLICANT: CHAWLA, Nandinder K.
APPLICANT: BAUGHN, Mariah R.
APPLICANT: MARQUIS, Joseph P.
APPLICANT: BECHA, Shanya D.
APPLICANT: KABLE, Amy E.
APPLICANT: LAL, Preeti G.
APPLICANT: RICHARDSON, Thomas W.
APPLICANT: LEE, Soo Yeun
APPLICANT: LEE, Ernestine A.
APPLICANT: TRAN, Bao
APPLICANT: WARREN, Bridget A.
APPLICANT: LU, Dying Aina M.
APPLICANT: GURURAJAN, Rajagopal
APPLICANT: SPRAGUE, William W.
APPLICANT: BLAKE, Julie J.
APPLICANT: THANGAVELU, Kavitha
APPLICANT: SWARNAKAR, Anita
APPLICANT: GORVAD, Ann E.
APPLICANT: GRIFFIN, Jennifer A.
APPLICANT: LINDQUIST, Erika A.
APPLICANT: ELLIOTT, Vicki S.
APPLICANT: ISON, Craig H.
APPLICANT: RAMUMAR, Jayalaxmi
TITLE OF INVENTION: MOLECULES FOR DISEASE DETECTION AND TREATMENT
FILE REFERENCE: PF-1232 PCT
CURRENT APPLICATION NUMBER: PCT/US02/32852
PRIORITY FILING DATE: 2002-10-10
PRIORITY FILING DATE: 2002-10-10
PRIORITY FILING DATE: 2001-10-12
PRIORITY FILING DATE: 2001-10-12
PRIORITY FILING DATE: 2001-10-26
PRIORITY FILING DATE: 2001-10-26
PRIORITY FILING DATE: 2001-11-02
PRIORITY FILING DATE: 2001-11-02
PRIORITY FILING DATE: 2001-11-09
PRIORITY FILING DATE: 2001-11-09
PRIORITY FILING DATE: 2001-11-16

NUMBER OF SEQ ID NOS: 96
SOFTWARE: PERL Program
SEQ ID NO 4
LENGTH: 270
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: Incyte ID No: 7494963CD1
PCT-US02-32852-4

Query Match
Best Local Similarity 55.4%; Score 31; DB 1; Length 270;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 DVEPTGAK 10
DB 80 DGMEPTGSR 88

RESULT 2
US-10-092-411A-3876
Sequence 3876, Application US/10092411A
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 032796-101
CURRENT APPLICATION NUMBER: US/10/092,411A
CURRENT FILING DATE: 2002-03-07
PRIOR APPLICATION NUMBER: US 09/134,001
PRIOR FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5676
SEQ ID NO 3876
LENGTH: 424
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-10-092-411A-3876

Query Match
Best Local Similarity 55.4%; Score 31; DB 6; Length 424;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 VEPTGA 9
DB 215 VEPTGA 220

RESULT 3
US-10-266-252-14
Sequence 14, Application US/10266252
GENERAL INFORMATION:
APPLICANT: Ghosh, Maladika
APPLICANT: Tang, Y. Tom
APPLICANT: Ren, Feiyan
TITLE OF INVENTION: Methods And Materials Relating To Stem Cell Growth Factor-Like
TITLE OF INVENTION: Polypeptides And Polynucleotides And Therapeutic Uses Thereof
FILE REFERENCE: HYS-44A
CURRENT APPLICATION NUMBER: US/10/266,252
CURRENT FILING DATE: 2002-10-07
PRIOR APPLICATION NUMBER: 60/327,731
PRIOR FILING DATE: 2001-10-05
PRIOR APPLICATION NUMBER: PCT/US00/35017
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 15

SOFTWARE: PatentIn version 3.1
SEQ ID NO 14
LENGTH: 435
TYPE: PRT
ORGANISM: homo sapiens
US-10-266-252-14

Query Match
Best Local Similarity 55.4%; Score 31; DB 6; Length 435;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 VEPTGAKE 11
DB 420 VEPTGEKE 427

RESULT 4
US-10-125-923A-472
Sequence 472, Application US/10125923A
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
TITLE OF INVENTION: ACIDS ENCODING THE SAME
FILE REFERENCE: P3430R1C79
CURRENT APPLICATION NUMBER: US/10/125,923A
CURRENT FILING DATE: 2002-01-15
PRIOR APPLICATION NUMBER: 10/052586
PRIOR FILING DATE: 2002-01-15
PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059266
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/063120
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063121
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063486
PRIOR FILING DATE: 1997-10-21
PRIOR APPLICATION NUMBER: 60/063540
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063541
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063544
PRIOR FILING DATE: 1997-10-28
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 472
LENGTH: 529
TYPE: PRT
ORGANISM: Homo Sapien
US-10-125-923A-472

Query Match
Best Local Similarity 55.4%; Score 31; DB 6; Length 529;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 VEPTGAKE 11
DB 514 VEPTGEKE 521

```
RESULT 5
US-09-555-534B-17
; Sequence 17, Application US/09555534B
; GENERAL INFORMATION:
; APPLICANT: Istituto Superiore di Sanit.
; APPLICANT: ENSOLI, Barbara
; TITLE OF INVENTION: HIV-1 Tat, or derivatives thereof, alone or in combination, for
; TITLE OF INVENTION: prophylactic and therapeutic vaccination against AIDS, tumors an
; FILE REFERENCE: 1354PTWO
; CURRENT APPLICATION NUMBER: US/09/555,534B
; PRIOR FILING DATE: 2002-11-05
; PRIOR APPLICATION NUMBER: RM97A000743
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 17
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Aids-associated retrovirus
US-09-555-534B-17

Query Match          51.8%; Score 29; DB 5; Length 14;
Best Local Similarity 71.4%; Pred. No. 4.4;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY      5 EPTGAK 11
      :||| 11
Db      8 DPTGPK 14

RESULT 6
US-09-555-534B-2
; Sequence 2, Application US/09555534B
; GENERAL INFORMATION:
; APPLICANT: Istituto Superiore di Sanit.
; APPLICANT: ENSOLI, Barbara
; TITLE OF INVENTION: HIV-1 Tat, or derivatives thereof, alone or in combination, for
; TITLE OF INVENTION: prophylactic and therapeutic vaccination against AIDS, tumors an
; FILE REFERENCE: 1354PTWO
; CURRENT APPLICATION NUMBER: US/09/555,534B
; CURRENT FILING DATE: 2002-11-05
; PRIOR APPLICATION NUMBER: RM97A000743
; PRIOR FILING DATE: 1997-12-01
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 2
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Aids-associated retrovirus
US-09-555-534B-2

Query Match          51.8%; Score 29; DB 5; Length 86;
Best Local Similarity 71.4%; Pred. No. 30;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY      5 EPTGAK 11
      :||| 11
Db      80 DPTGPK 86

RESULT 7
US-09-555-534B-4
; Sequence 4, Application US/09555534B
; GENERAL INFORMATION:
; APPLICANT: Istituto Superiore di Sanit.
; APPLICANT: ENSOLI, Barbara
; TITLE OF INVENTION: HIV-1 Tat, or derivatives thereof, alone or in combination, for
; TITLE OF INVENTION: prophylactic and therapeutic vaccination against AIDS, tumors an
; FILE REFERENCE: 1354PTWO
; CURRENT APPLICATION NUMBER: US/09/555,534B

; Sequence 6, Application US/09555534B
; GENERAL INFORMATION:
; APPLICANT: Istituto Superiore di Sanit.
; APPLICANT: ENSOLI, Barbara
; TITLE OF INVENTION: HIV-1 Tat, or derivatives thereof, alone or in combination, fo
; TITLE OF INVENTION: prophylactic and therapeutic vaccination against AIDS, tumors
; FILE REFERENCE: 1354PTWO
; CURRENT APPLICATION NUMBER: US/09/555,534B
; CURRENT FILING DATE: 2002-11-05
; PRIOR APPLICATION NUMBER: RM97A000743
; PRIOR FILING DATE: 1997-12-01
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 6
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Aids-associated retrovirus
US-09-555-534B-6

Query Match          51.8%; Score 29; DB 5; Length 86;
Best Local Similarity 71.4%; Pred. No. 30;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY      5 EPTGAK 11
      :||| 11
Db      80 DPTGPK 86

RESULT 8
US-09-555-534B-6
; Sequence 6, Application US/09555534B
; GENERAL INFORMATION:
; APPLICANT: Istituto Superiore di Sanit.
; APPLICANT: ENSOLI, Barbara
; TITLE OF INVENTION: HIV-1 Tat, or derivatives thereof, alone or in combination, fo
; TITLE OF INVENTION: prophylactic and therapeutic vaccination against AIDS, tumors
; FILE REFERENCE: 1354PTWO
; CURRENT APPLICATION NUMBER: US/09/555,534B
; CURRENT FILING DATE: 2002-11-05
; PRIOR APPLICATION NUMBER: RM97A000743
; PRIOR FILING DATE: 1997-12-01
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 6
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Aids-associated retrovirus
US-09-555-534B-6

Query Match          51.8%; Score 29; DB 5; Length 86;
Best Local Similarity 71.4%; Pred. No. 30;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY      5 EPTGAK 11
      :||| 11
Db      80 DPTGPK 86

RESULT 9
PCT-US02-00215-8
; Sequence 8, Application PC/TUS0200215
; GENERAL INFORMATION:
; APPLICANT: Meng, Xiang-Jin
; APPLICANT: Haqshenas, Gholamreza
; APPLICANT: Huang, Fang-Fang
; TITLE OF INVENTION: AVIAN HEPATITIS E VIRUS, VACCINES AND METHODS OF PROTECTING
; TITLE OF INVENTION: AGAINST AVIAN HEPATITIS-SPLENOMEGALY SYNDROME AND MAMMALIAN
; FILE REFERENCE: AM100389 PCT
; CURRENT APPLICATION NUMBER: PCT/US02/00215
; CURRENT FILING DATE: 2002-10-07
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 8
; LENGTH: 87
; TYPE: PRT
; ORGANISM: Hepatitis E virus
PCT-US02-00215-8

Query Match          51.8%; Score 29; DB 1; Length 87;
```

Best Local Similarity 55.6%; Pred. No. 30;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 ADVEPTGA 9
| : : : |
Db 51 AVTI0PSGA 59

RESULT 10
US-10-092-411A-3835

; Sequence 3835, Application US/10092411A
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: 032796-101
; CURRENT APPLICATION NUMBER: US/10/092,411A
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 09/134,001
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5676
; SEQ ID NO 3835
; LENGTH: 148
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-10-092-411A-3835

Query Match 51.8%; Score 29; DB 6; Length 148;
Best Local Similarity 55.6%; Pred. No. 53;

Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 DIVEPTGAK 10
| : : : |
Db 61 DASQPTGAQ 69

RESULT 11

US-09-905-666A-76
; Sequence 76, Application US/09905666A
; GENERAL INFORMATION:
; APPLICANT: GIVER, LORRAINE J.
; APPLICANT: MINSHULL, JEREMY
; APPLICANT: VOGEL, KURT
; TITLE OF INVENTION: NOVEL LIPASE GENES
; FILE REFERENCE: 0184.310US
; CURRENT APPLICATION NUMBER: US/09/905,666A
; CURRENT FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: 60/217,954
; PRIOR FILING DATE: 2000-07-13
; PRIOR APPLICATION NUMBER: 60/300,378
; PRIOR FILING DATE: 2001-06-21
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 76
; LENGTH: 180
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
US-09-905-666A-76

Query Match 51.8%; Score 29; DB 5; Length 180;
Best Local Similarity 50.0%; Pred. No. 65;

Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 DIVEPTGAK 11
| : : : |
Db 60 DVLDTGAKK 69

RESULT 12
US-09-905-666A-64

; Sequence 64, Application US/09905666A
; GENERAL INFORMATION:
; APPLICANT: GIVER, LORRAINE J.
; APPLICANT: MINSHULL, JEREMY
; APPLICANT: VOGEL, KURT
; TITLE OF INVENTION: NOVEL LIPASE GENES
; FILE REFERENCE: 0184.310US
; CURRENT APPLICATION NUMBER: US/09/905,666A
; CURRENT FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: 60/217,954
; PRIOR FILING DATE: 2000-07-13
; PRIOR APPLICATION NUMBER: 60/300,378
; PRIOR FILING DATE: 2001-06-21
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 64
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Bacillus sp.
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (73)
; OTHER INFORMATION: Variable amino acid
US-09-905-666A-64

Query Match 51.8%; Score 29; DB 5; Length 212;
Best Local Similarity 50.0%; Pred. No. 77;

Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 DIVEPTGAK 11
| : : : |
Db 92 DVLDTGAKK 101

RESULT 13

US-09-905-666A-71
; Sequence 71, Application US/09905666A
; GENERAL INFORMATION:
; APPLICANT: GIVER, LORRAINE J.
; APPLICANT: MINSHULL, JEREMY
; APPLICANT: VOGEL, KURT
; TITLE OF INVENTION: NOVEL LIPASE GENES
; FILE REFERENCE: 0184.310US
; CURRENT APPLICATION NUMBER: US/09/905,666A
; CURRENT FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: 60/217,954
; PRIOR FILING DATE: 2000-07-13
; PRIOR APPLICATION NUMBER: 60/300,378
; PRIOR FILING DATE: 2001-06-21
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 71
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-905-666A-71

Query Match 51.8%; Score 29; DB 5; Length 212;
Best Local Similarity 50.0%; Pred. No. 77;

Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 DIVEPTGAK 11
| : : : |
Db 92 DVLDTGAKK 101

RESULT 14

US-09-905-666A-72
; Sequence 72, Application US/09905666A

```
; GENERAL INFORMATION:
; APPLICANT: GIVER, LORRAINE J.
; APPLICANT: MINSHULL, JEREMY
; APPLICANT: VOGEL, KURT
; TITLE OF INVENTION: NOVEL LIPASE GENES
; FILE REFERENCE: 0184.310US
; CURRENT APPLICATION NUMBER: US/09/905,666A
; CURRENT FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: 60/217,954
; PRIOR FILING DATE: 2000-07-13
; PRIOR APPLICATION NUMBER: 60/300,378
; PRIOR FILING DATE: 2001-06-21
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 72
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-905-666A-72
```

```
Query Match          51.8%; Score 29; DB 5; Length 212;
Best Local Similarity 50.0%; Pred. No. 77;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
```

```
OY 2 DIVEPTGAK 11
1 : : ||||:
DB 92 DVLDKTGAKK 101
```

```
RESULT 15
US-09-905-666A-74
; Sequence 74, Application US/09905666A
; GENERAL INFORMATION:
; APPLICANT: GIVER, LORRAINE J.
; APPLICANT: MINSHULL, JEREMY
; APPLICANT: VOGEL, KURT
; TITLE OF INVENTION: NOVEL LIPASE GENES
; FILE REFERENCE: 0184.310US
; CURRENT APPLICATION NUMBER: US/09/905,666A
; CURRENT FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: 60/217,954
; PRIOR FILING DATE: 2000-07-13
; PRIOR APPLICATION NUMBER: 60/300,378
; PRIOR FILING DATE: 2001-06-21
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 74
; LENGTH: 213
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-905-666A-74
```

```
Query Match          51.8%; Score 29; DB 5; Length 213;
Best Local Similarity 50.0%; Pred. No. 78;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
```

```
OY 2 DIVEPTGAK 11
1 : : ||||:
DB 92 DVLDKTGAKK 101
```

Search completed: November 18, 2002, 16:05:21
Job time : 3.90323 secs

THIS PAGE BLANK (USP10)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 17:21:21 : Search time 3223 Seconds
(without alignments) 234.029 Million cell updates/sec

Title: US-09-848-781-4

Perfect score: 30
Sequence: 1 catgcctggtgtagaacgycgtgataccca 30

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 24791104 seqs, 12571243825 residues

Total number of hits satisfying chosen parameters: 49582208

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

```
Pending_Patents_NA_Main:*
1: /cgn2_6/ptodata/1/pna/US06000_COMB.seq:*
2: /cgn2_6/ptodata/1/pna/US06_COMB.seq:*
3: /cgn2_6/ptodata/1/pna/US07_COMB.seq:*
4: /cgn2_6/ptodata/1/pna/US080_COMB.seq:*
5: /cgn2_6/ptodata/1/pna/US081_COMB.seq:*
6: /cgn2_6/ptodata/1/pna/US082_COMB.seq:*
7: /cgn2_6/ptodata/1/pna/US083_COMB.seq:*
8: /cgn2_6/ptodata/1/pna/US084_COMB.seq:*
9: /cgn2_6/ptodata/1/pna/US085_COMB.seq:*
10: /cgn2_6/ptodata/1/pna/US086_COMB.seq:*
11: /cgn2_6/ptodata/1/pna/US087_COMB.seq:*
12: /cgn2_6/ptodata/1/pna/US088_COMB.seq:*
13: /cgn2_6/ptodata/1/pna/US089_COMB.seq:*
14: /cgn2_6/ptodata/1/pna/US090_COMB.seq:*
15: /cgn2_6/ptodata/1/pna/US091_COMB.seq:*
16: /cgn2_6/ptodata/1/pna/US092_COMB.seq:*
17: /cgn2_6/ptodata/1/pna/US093_COMB.seq:*
18: /cgn2_6/ptodata/1/pna/US094_COMB.seq:*
19: /cgn2_6/ptodata/1/pna/US095_COMB.seq:*
20: /cgn2_6/ptodata/1/pna/US095B_COMB.seq:*
21: /cgn2_6/ptodata/1/pna/US095C_COMB.seq:*
22: /cgn2_6/ptodata/1/pna/US095D_COMB.seq:*
23: /cgn2_6/ptodata/1/pna/US096B_COMB.seq:*
24: /cgn2_6/ptodata/1/pna/US096C_COMB.seq:*
25: /cgn2_6/ptodata/1/pna/US096D_COMB.seq:*
26: /cgn2_6/ptodata/1/pna/US096E_COMB.seq:*
27: /cgn2_6/ptodata/1/pna/US096F_COMB.seq:*
28: /cgn2_6/ptodata/1/pna/US097A_COMB.seq:*
29: /cgn2_6/ptodata/1/pna/US097B_COMB.seq:*
30: /cgn2_6/ptodata/1/pna/US097C_COMB.seq:*
31: /cgn2_6/ptodata/1/pna/US098A_COMB.seq:*
32: /cgn2_6/ptodata/1/pna/US098B_COMB.seq:*
33: /cgn2_6/ptodata/1/pna/US098C_COMB.seq:*
34: /cgn2_6/ptodata/1/pna/US099A_COMB.seq:*
35: /cgn2_6/ptodata/1/pna/US099B_COMB.seq:*
36: /cgn2_6/ptodata/1/pna/US099C_COMB.seq:*
37: /cgn2_6/ptodata/1/pna/US099D_COMB.seq:*
38: /cgn2_6/ptodata/1/pna/US100A_COMB.seq:*
39: /cgn2_6/ptodata/1/pna/US100B_COMB.seq:*
40: /cgn2_6/ptodata/1/pna/US101A_COMB.seq:*
41: /cgn2_6/ptodata/1/pna/US101B_COMB.seq:*
42: /cgn2_6/ptodata/1/pna/US102A_COMB.seq:*
43: /cgn2_6/ptodata/1/pna/US102B_COMB.seq:*
```

Result No.	Score	Query Match	Length	ID	Description
1	30	100.0	30	US-09-848-781-4	Sequence 4, Appl 1
2	20.4	68.0	484	US-08-803-610C-2232	Sequence 2232, Ap
3	20.4	68.0	484	US-08-803-610D-2232	Sequence 2232, Ap
4	20.4	68.0	484	US-08-803-610E-2232	Sequence 2232, Ap
5	19.4	64.7	1609	US-09-634-306B-263871	Sequence 263871,
6	19.4	64.7	1609	US-09-634-306B-263872	Sequence 263872,
7	19.4	64.7	1609	US-09-634-306B-263873	Sequence 263873,
8	19.4	64.7	1609	US-10-027-632-263871	Sequence 263871,
9	19.4	64.7	1609	US-10-027-632-263872	Sequence 263872,
10	19.4	64.7	1609	US-10-027-632-263873	Sequence 263873,
11	19.2	64.0	3108	US-09-614-221A-194	Sequence 194, App
12	19.2	64.0	3108	US-09-360-039-45594	Sequence 45594, A
13	19	63.3	180	US-09-565-240-21674	Sequence 21674, A
14	19	63.3	180	US-09-654-617-71858	Sequence 71858, A
15	19	63.3	180	US-09-684-016-71858	Sequence 71858, A
16	18.8	62.7	284	US-09-436-762A-9978	Sequence 9978, Ap
17	18.8	62.7	295	US-09-540-764-56898	Sequence 56898, A
18	18.8	62.7	317	US-09-394-745-9938	Sequence 9938, A
19	18.8	62.7	317	US-09-565-306-31890	Sequence 31890, A
20	18.8	62.7	317	US-09-654-617-318101	Sequence 318101,
21	18.8	62.7	317	US-09-684-016-318101	Sequence 318101,

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

```
22 18.8 62.7 417 36 US-09-963-299-906 Sequence 906, App
23 18.8 62.7 463 17 US-09-395-545-1555 Sequence 1555, Ap
24 18.8 62.7 463 31 US-09-825-774-1555 Sequence 1555, Ap
25 18.8 62.7 526 62 US-60-180-489-6155 Sequence 6155, Ap
26 18.8 62.7 543 63 US-60-195-135-1877 Sequence 1877, App
27 18.8 62.7 3104 24 US-09-634-306B-111724 Sequence 111724,
28 18.8 62.7 3104 24 US-09-634-306B-115782 Sequence 115782,
29 18.8 62.7 3104 24 US-09-634-306B-115783 Sequence 115783,
30 18.8 62.7 3104 38 US-10-027-632-111724 Sequence 111724,
31 18.8 62.7 3104 38 US-10-027-632-115782 Sequence 115782,
32 18.8 62.7 3104 38 US-10-027-632-115783 Sequence 115783,
33 18.8 62.7 194788 19 US-09-528-237A-1930 Sequence 1930, Ap
34 18.4 61.3 159 29 US-09-724-750-8119 Sequence 8119, Ap
35 18.4 61.3 159 61 US-60-171-431-8119 Sequence 8119, Ap
36 18.4 61.3 241 69 US-60-253-652-7947 Sequence 7947, Ap
37 18.4 61.3 358 57 US-60-132-861-905 Sequence 905, App
38 18.4 61.3 452 57 US-60-138-103-7159 Sequence 7159, Ap
39 18.4 61.3 503 57 US-60-132-861-1577 Sequence 1577, Ap
40 18.4 61.3 503 57 US-60-132-861-2077 Sequence 2077, Ap
41 18.4 61.3 570 33 US-09-864-761-16234 Sequence 16234, A
42 18.4 61.3 570 42 US-10-203-134-12975 Sequence 12975, A
43 18.4 61.3 577 64 US-60-207-458-28910 Sequence 28910, A
44 18.4 61.3 642 56 US-60-126-265-3719 Sequence 3719, Ap
45 18.4 61.3 687 16 US-09-252-691-1298 Sequence 1298, Ap
```

ALIGNMENTS

```
RESULT 1
US-09-848-781-4
; Sequence 4, Application US/09848781
; GENERAL INFORMATION:
; APPLICANT: SANDERS, MITCHELL
; TITLE OF INVENTION: A DEVICE FOR DETECTING BACTERIAL CONTAMINATION AND METHOD OF USE
; FILE REFERENCE: 102951-10
; CURRENT APPLICATION NUMBER: US/09/848, 781
; CURRENT FILING DATE: 2001-05-03
; PRIOR APPLICATION NUMBER: 60/201,405
; PRIOR FILING DATE: 2000-05-03
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 4
; LENGTH: 30
; TYPE: DNA
; ORGANISM: Listeria monocytogenes
US-09-848-781-4

Query Match 100.0%; Score 30; DB 32; Length 30;
Best Local Similarity 100.0%; Pred. No. 0.004;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATGCCATGGGTAGAACGGGCTGATACCA 30
Db 1 CATGCCATGGGTAGAACGGGCTGATACCA 30

RESULT 2
US-08-803-610C-2232
; Sequence 2232, Application US/08803610C
; GENERAL INFORMATION:
; APPLICANT: Craig A. Rosen
; APPLICANT: Patrick S. Dillon
; APPLICANT: Julie Earle-Hughes
; APPLICANT: Haodong Li
; APPLICANT: Steven M. Ruben
; APPLICANT: William A. Haseltine
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products - P024
; NUMBER OF SEQUENCES: 13764
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
```

```
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/803,610C
; FILING DATE: Feb-21-97
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/011,991
; FILING DATE: Feb-21-96
; ATTORNEY/AGENT INFORMATION:
; NAME: A. Anders Brookes
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: P024
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 2232:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 484 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-803-610C-2232
```

```
Query Match 68.0%; Score 20.4; DB 12; Length 484;
Best Local Similarity 80.0%; Pred. No. 1.3e+02;
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
```

```
QY 1 CATGCCGTGGGAGACGCTGATACCA 30
Db 123 CATGCCGTGGGAGACGCTGATACCA 152

RESULT 3
US-08-803-610D-2232
; Sequence 2232, Application US/08803610D
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig, et al.
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products-24
; FILE REFERENCE: PO-24
; CURRENT APPLICATION NUMBER: US/08/803,610D
; CURRENT FILING DATE: 1997-02-21
; PRIOR APPLICATION NUMBER: 60/011,991
; PRIOR FILING DATE: 1996-02-21
; NUMBER OF SEQ ID NOS: 13764
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2232
; LENGTH: 484
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (108)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (221)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (257)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (264)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (288)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
```

```
LOCATION: (357)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (384)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (414)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (424)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (432)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (468)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (474)
OTHER INFORMATION: n equals a,t,g, or c
US-08-803-610D-2232
```

Query Match 68.0%; Score 20.4; DB 12; Length 484;
Best Local Similarity 80.0%; Pred. No. 1.3e+02;

Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```
OY 1 CATGCCATGGTGAACGGGCTGATACCCA 30
Db 123 CATGCCGTGGGAACGACGTGATCCCA 152
```

RESULT 4
US-08-803-610E-2232

Sequence 2232, Application US/08803610E

GENERAL INFORMATION:

APPLICANT: Rosen, Craig, et al.

TITLE OF INVENTION: Human Genes, Sequences, and Expression Products-24

FILE REFERENCE: PO-24

CURRENT APPLICATION NUMBER: US/08/803,610E

PRIOR FILING DATE: 1997-02-21

PRIOR APPLICATION NUMBER: 60/011,991

PRIOR FILING DATE: 1996-02-21

NUMBER OF SEQ ID NOS: 13764

SOFTWARE: Patentln Ver. 2.0

SEQ ID NO 2232

LENGTH: 484

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc feature

LOCATION: (108)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (221)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (257)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (264)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (288)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (357)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (384)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (414)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (414)

OTHER INFORMATION: n equals a,t,g, or c

```
NAME/KEY: misc feature
LOCATION: (424)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (432)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (468)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (474)
OTHER INFORMATION: n equals a,t,g, or c
US-08-803-610E-2232
```

Query Match 68.0%; Score 20.4; DB 12; Length 484;
Best Local Similarity 80.0%; Pred. No. 1.3e+02;

Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```
OY 1 CATGCCATGGTGAACGGGCTGATACCCA 30
Db 123 CATGCCGTGGGAACGACGTGATCCCA 152
```

RESULT 5

US-09-634-306B-263871

Sequence 263871, Application US/09634306B

GENERAL INFORMATION:

APPLICANT: Wang, David G.

TITLE OF INVENTION: Identification and Mapping of Single Nucleotide

FILE REFERENCE: 108827.129

CURRENT APPLICATION NUMBER: US/09/634,306B

PRIOR FILING DATE: 2002-02-21

PRIOR APPLICATION NUMBER: US 60/218,006

PRIOR FILING DATE: 2000-07-12

PRIOR APPLICATION NUMBER: US 60/198,676

PRIOR FILING DATE: 2000-04-20

PRIOR APPLICATION NUMBER: US 60/193,483

PRIOR FILING DATE: 2000-03-29

PRIOR APPLICATION NUMBER: US 60/185,218

PRIOR FILING DATE: 2000-02-24

PRIOR APPLICATION NUMBER: US 60/167,363

PRIOR FILING DATE: 1999-11-23

PRIOR APPLICATION NUMBER: US 60/156,358

PRIOR FILING DATE: 1999-09-28

PRIOR APPLICATION NUMBER: US 60/146,002

PRIOR FILING DATE: 1999-08-09

NUMBER OF SEQ ID NOS: 325720

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 263871

LENGTH: 1609

TYPE: DNA

ORGANISM: Human

US-09-634-306B-263871

Query Match 64.7%; Score 19.4; DB 24; Length 1609;
Best Local Similarity 79.3%; Pred. No. 4.2e+02;

Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```
OY 2 ATGCCATGGTGAACGGGCTGATACCCA 30
Db 1242 ATGTAATGGGTGAAGGATCTACCCA 1270
```

RESULT 6

US-09-634-306B-263872

Sequence 263872, Application US/09634306B

GENERAL INFORMATION:

APPLICANT: Wang, David G.

TITLE OF INVENTION: Identification and Mapping of Single Nucleotide

FILE REFERENCE: 108827.129

CURRENT APPLICATION NUMBER: US/09/634,306B

```

; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 263872
; LENGTH: 1609
; TYPE: DNA
; ORGANISM: Human
US-09-634-306B-263872
```

```
Query Match
Best Local Similarity 64.7%; Score 19.4; DB 24; Length 1609;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
```

```
OY 2 ATGCATGGGTAGACGGGCTGATACCA 30
DB 1242 ATGTATGGGTAGTAAAGGATGCTACCA 1270
```

```

RESULT 7
US-09-634-306B-263873
; Sequence 263873, Application US/09634306B
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/09/634,306B
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 263873
; LENGTH: 1609
; TYPE: DNA
; ORGANISM: Human
US-09-634-306B-263873
```

```
Query Match
Best Local Similarity 64.7%; Score 19.4; DB 24; Length 1609;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
```

```
OY 2 ATGCATGGGTAGACGGGCTGATACCA 30
DB 1242 ATGTATGGGTAGTAAAGGATGCTACCA 1270
```

```

RESULT 8
US-10-027-632-263871
; Sequence 263871, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 263871
; LENGTH: 1609
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-263871
```

```
Query Match
Best Local Similarity 64.7%; Score 19.4; DB 38; Length 1609;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
```

```
OY 2 ATGCATGGGTAGACGGGCTGATACCA 30
DB 1242 ATGTATGGGTAGTAAAGGATGCTACCA 1270
```

```

RESULT 9
US-10-027-632-263872
; Sequence 263872, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 263872
; LENGTH: 1609
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-263872
```

Query Match 64.7%; Score 19.4; DB 38; Length 1609;
Best Local Similarity 79.3%; Pred. No. 4.2e+02;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 ATGCCATGGGTAGAACGGGCTGATACCA 30
DB 1242 ATGTAATGGGTAGTAAGGATGCTACCA 1270

RESULT 10

US-10-027-632-263873
; Sequence 263873, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT FILING DATE: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 263873
; LENGTH: 1609
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-263873

Query Match 64.7%; Score 19.4; DB 38; Length 1609;
Best Local Similarity 79.3%; Pred. No. 4.2e+02;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 ATGCCATGGGTAGAACGGGCTGATACCA 30
DB 1242 ATGTAATGGGTAGTAAGGATGCTACCA 1270

RESULT 11

US-09-614-221A-194/c
; Sequence 194, Application US/09614221A
; GENERAL INFORMATION:
; APPLICANT: Karunanandaa, Balasubramanian
; APPLICANT: Yu, Jaehyuk
; APPLICANT: Kishore, Ganesh M.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED
; FILE REFERENCE: 16516.075
; CURRENT FILING DATE: US/09/614,221A
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/142,981
; PRIOR FILING DATE: 1999-07-12
; NUMBER OF SEQ ID NOS: 626
; SEQ ID NO 194
; LENGTH: 3108
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-614-221A-194

Query Match 64.0%; Score 19.2; DB 23; Length 3108;
Best Local Similarity 87.5%; Pred. No. 5.5e+02;

Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 6 CATGGGTAGAACGGGCTGATACCC 29
DB 832 CATGGGTAGAACGGGCTGATACCC 809

RESULT 12

US-60-360-039-45594/c
; Sequence 45594, Application US/60360039
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Chen, Xianfeng
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)A
; CURRENT FILING DATE: US/60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 45594
; LENGTH: 3108
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-60-360-039-45594

Query Match 64.0%; Score 19.2; DB 80; Length 3108;
Best Local Similarity 87.5%; Pred. No. 5.5e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 6 CATGGGTAGAACGGGCTGATACCC 29
DB 832 CATGGGTAGAACGGGCTGATACCC 809

RESULT 13

US-09-565-240-21674
; Sequence 21674, Application US/09565240
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S.
; APPLICANT: La Rosa, Thomas J.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated w/
; FILE REFERENCE: 38-21(15793)B
; CURRENT FILING DATE: US/09/565,240
; PRIOR FILING DATE: 2000-05-03
; NUMBER OF SEQ ID NOS: 54005
; SEQ ID NO 21674
; LENGTH: 180
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: LIB3106-048-Q1-K1-A5
US-09-565-240-21674

Query Match 63.3%; Score 19; DB 22; Length 180;
Best Local Similarity 81.5%; Pred. No. 5e+02;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 GCCATGGGTAGAACGGGCTGATACCA 30
DB 82 GCCATGGGTAGAACGGGCTGATACCA 108

RESULT 14

US-09-654-617-71858
; Sequence 71858, Application US/09654617
; GENERAL INFORMATION:
; APPLICANT: Kovall, David K.
; APPLICANT: Liu, Jingdong
; TITLE OF INVENTION: Annotated Plant Genes
; FILE REFERENCE: 38-21(15097)D

```

; CURRENT APPLICATION NUMBER: US/09/654,617
; CURRENT FILING DATE: 2000-09-05
; NUMBER OF SEQ ID NOS: 463173
; SEQ ID NO 71858
; LENGTH: 180
; TYPE: DNA
; ORGANISM: Glycine max
US-09-654-617-71858

```

```

Query Match      63.3%; Score 19; DB 25; Length 180;
Best Local Similarity 81.5%; Pred. No. 5e+02;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

```

QY 4 GCCATGGGTAGAACGGGCTGTATACCA 30
    |||| ||||| ||||| ||||
DB 82 GCCAGAGTAGAAGGCGCTGTGCCCCA 108

```

```

RESULT 15
US-09-684-016-71858
; Sequence 71858, Application US/09684016
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; TITLE OF INVENTION: Annotated Plant Genes
; FILE REFERENCE: 38-21(15097)D
; CURRENT APPLICATION NUMBER: US/09/684,016
; CURRENT FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: US 09/654,617
; PRIORITY FILING DATE: 2000-09-05
; NUMBER OF SEQ ID NOS: 463173
; SEQ ID NO 71858
; LENGTH: 180
; TYPE: DNA
; ORGANISM: Glycine max
US-09-684-016-71858

```

```

Query Match      63.3%; Score 19; DB 27; Length 180;
Best Local Similarity 81.5%; Pred. No. 5e+02;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

```

QY 4 GCCATGGGTAGAACGGGCTGTATACCA 30
    |||| ||||| ||||| ||||
DB 82 GCCAGAGTAGAAGGCGCTGTGCCCCA 108

```

Search completed: November 23, 2002, 19:39:01
 Job time : 3232 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 17:34:36 ; Search time 53 Seconds
(without alignments)
243.500 Million cell updates/sec

Title: US-09-848-781-4

Perfect score: 30

Sequence: 1 catgcacatggttagaacgctgataccca 30

Scoring table: IDENTITY_NNC

Searched: 280232 seqs, 215092063 residues

Total number of hits satisfying chosen parameters: 560464

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Pending Patents_NA_New:*
1: /cgn2_6/ptodata/2/pna/US05_NEW_COMB.seq:*
2: /cgn2_6/ptodata/2/pna/US06_NEW_COMB.seq:*
3: /cgn2_6/ptodata/2/pna/US07_NEW_COMB.seq:*
4: /cgn2_6/ptodata/2/pna/US08_NEW_COMB.seq:*
5: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:*
6: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq:*
7: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	17.2	57.3	627	US-10-109-368-31	Sequence 31, Appl
2	17.2	57.3	35829	PCT-US02-32727-68	Sequence 68, Appl
3	17	56.7	1773	PCT-US02-05068-155	Sequence 15, Appl
4	17	56.7	21784	US-10-274-031-3	Sequence 3, Appl1
5	16.6	55.3	750	US-09-724-676-27540	Sequence 27540, A
6	16.6	55.3	1074	US-10-264-237-1263	Sequence 1263, Ap
7	16.6	54.7	1422	US-09-724-676-27541	Sequence 27541, A
8	16.4	54.7	2064	US-10-274-694-27	Sequence 27, Appl
9	16.4	54.7	2626	US-10-152-319A-1686	Sequence 1686, Ap
10	16.4	54.7	8924	US-10-264-237-2850	Sequence 2850, Ap
11	16.4	54.7	8925	US-10-264-237-2849	Sequence 2849, Ap
12	16.4	54.7	13825	US-10-264-213-27	Sequence 27, Appl
13	16.2	54.0	192	US-09-513-999C-14199	Sequence 14199, A
14	16.2	54.0	1110	US-09-724-676-29903	Sequence 29903, A
15	16.2	54.0	1124	US-09-724-676-29902	Sequence 29902, A
16	16.2	54.0	1588	US-09-724-676-29902	Sequence 29902, A
17	16.2	54.0	1602	US-09-724-676-29889	Sequence 29889, A
18	16.2	54.0	2390	US-09-724-676-29901	Sequence 29901, A
19	16.2	54.0	2404	US-09-724-676-29901	Sequence 29898, A
20	16.2	54.0	4316	US-10-257-021-81	Sequence 81, Appl
21	16.2	54.0	4851	US-09-724-676-29907	Sequence 29907, A
22	16.2	54.0	5329	US-09-724-676-29906	Sequence 29906, A
23	16.2	54.0	6131	US-09-724-676-29905	Sequence 29905, A
24	16.2	54.0	66788	PCT-US02-32727-10	Sequence 10, Appl
25	16	53.3	673	US-10-240-425-65	Sequence 65, Appl
26	16	53.3	1204	PCT-US02-32727-266	Sequence 266, Appl

c	27	16	53.3	29559	1	PCT-US02-32727-41	Sequence 41, Appl
c	28	15.8	52.7	296	5	US-09-513-999C-26946	Sequence 26946, A
c	29	15.8	52.7	877	5	US-09-724-676-20037	Sequence 20037, A
c	30	15.8	52.7	1042	5	US-09-724-676-20036	Sequence 20036, A
c	31	15.8	52.7	1080	6	US-10-173-143-9	Sequence 9, Appl1
c	32	15.8	52.7	1101	5	US-09-724-676-20035	Sequence 20035, A
c	33	15.8	52.7	1726	5	US-09-978-418-23	Sequence 23, Appl
c	34	15.8	52.7	2063	6	US-10-240-425-1565	Sequence 1565, Ap
c	35	15.8	52.7	2237	6	US-10-152-319A-1603	Sequence 1603, Ap
c	36	15.8	52.7	2620	6	US-10-152-319A-2123	Sequence 2123, Ap
c	37	15.8	52.7	2868	1	PCT-US02-32727-163	Sequence 163, Appl
c	38	15.8	52.7	19717	1	PCT-US02-32727-33	Sequence 33, Appl
c	39	15.8	52.7	21263	1	PCT-US02-32727-107	Sequence 107, Appl
c	40	15.8	52.7	56050	1	PCT-US02-32727-44	Sequence 44, Appl
c	41	15.8	52.7	201143	6	US-10-240-425-1099	Sequence 1099, Ap
c	42	15.8	52.7	276820	6	PCT-US02-32727-9	Sequence 9, Appl1
c	43	15.8	52.0	339	5	US-10-271-416-9	Sequence 339, A
c	44	15.6	52.0	339	5	US-09-513-999C-36369	Sequence 36369, A
c	45	15.6	52.0	359	5	US-09-620-607B-869	Sequence 869, Appl

ALIGNMENTS

RESULT 1

US-10-109-368-31/C
Sequence 31, Application US/10109368

GENERAL INFORMATION:

APPLICANT: Miles, Vincent J.
Mathews, Michael B.
Katz, Michael G.
Witherell, Gary
Watson, Julia C.

TITLE OF INVENTION: METHOD FOR SELECTIVE INACTIVATION
OF VIRAL REPLICATION

NUMBER OF SEQUENCE: 33

CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036/2711

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/109,368

FILING DATE: 27-Mar-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/221,816

FILING DATE: 01-Apr-1994

ATTORNEY/AGENT INFORMATION:

NAME: Coruzzi, Laura A

REGISTRATION NUMBER: 30,742

REFERENCE/DOCKET NUMBER: 7960-030

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 790-9090

TELEFAX: (212) 869-8864

TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 31:

SEQUENCE CHARACTERISTICS:

LENGTH: 627 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: RNA

SEQUENCE DESCRIPTION: SEQ ID NO: 31:

US-10-109-368-31

	Query Match	Similarity	Score	DB	Length
Best	Local	73.3%	17.2	6	627
Matches	22	Conservative	0	Mismatches	8
				Indels	0
				Gaps	0

```

RESULT 2
PCT-US02-32727-68
: Sequence 68, Application PC/TUS0232727
: GENERAL INFORMATION:
: APPLICANT: Mitcham, Jennifer
: APPLICANT: Skeiky, Yasir
: APPLICANT: Persing, David
: APPLICANT: Bhatia, Ajay
: APPLICANT: Maisonneuve, Jean Francois
: APPLICANT: Zhang, Yanni
: APPLICANT: Wang, Siqing
: APPLICANT: Jen, Shyian
: APPLICANT: Lodes, Michael
: APPLICANT: Benson, Darin
: APPLICANT: Jones, Robert
: APPLICANT: Carter, Darlick
: APPLICANT: Barth, Brenda
: APPLICANT: Douglass, John
: TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acnes V
: FILE REFERENCE: 210121.514C1
: CURRENT APPLICATION NUMBER: PCT/US02/32727
: CURRENT FILING DATE: 2002-10-11
: NUMBER OF SEQ ID NOS: 30992
: SEQ ID NO 68
: LENGTH: 35829
: TYPE: DNA
: ORGANISM: Propionibacterium acnes
PCT-US02-32727-68

```

Query Match	57.3%	Score 17.2;	DB 1;	Length 35829;
Best Local Similarity	73.3%	Pred. No. 1.1e+02;		
Matches 22; Conservative	0;	Mismatches 8;	Indels 0;	Gaps 0;

Qy 1 CATGCCATGGGTAGAACGGGGCTGATACCCA 30
||| || |||| ||||| |||
Db 24751 CATCCCCGGGGTCTAACGGGGCTGACGCCGA 24780

```

RESULT 3
PCT-US02-05068-155
Sequence 155, Application PC/7050205068
GENERAL INFORMATION:
APPLICANT: Diversa Corporation
TITLE OF INVENTION: ENZYMES HAVING ALPHA AMYLASE ACTIVITY
TITLE OF INVENTION: ENZYMES HAVING ALPHA AMYLASE ACTIVITY
FILE REFERENCE: 09010-108W01
CURRENT APPLICATION NUMBER: PCT/US02/05068
CURRENT FILING DATE: 2002-02-21
PRIORITY APPLICATION NUMBER: US 60/270,495
PRIORITY FILING DATE: 2001-02-21
PRIORITY APPLICATION NUMBER: US 60/270,496
PRIORITY FILING DATE: 2001-02-21
PRIORITY APPLICATION NUMBER: US 60/291,122
PRIORITY FILING DATE: 2001-05-14
NUMBER OF SEQ ID NOS: 321
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 155
LENGTH: 1773
TYPE: DNA
ORGANISM: Environmental
PCT-US02-05068-155

```

Query Match	56.78;	Score 17;	DB 1;	Length 1773;
Best Local Similarity	80.0%;	Pred. No. 78;		

	Matches	20;	Conservative	0;	Mismatches	5;	Indels	0;	Gaps	0;
OY	3	TGCCATGGGTAGACGGCTGATAC	27							
Db	917	TGCACATGGGTACATTCCCATGTATAC	941							

```

RESULT 4
US-10-274-031-3/C
; Sequence 3, Application US/10274031
; GENERAL INFORMATION:
; APPLICANT: Gan, weiniu et al.
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: C1001194DIY
; CURRENT APPLICATION NUMBER: US/10/274,031
; CURRENT FILING DATE: 2002-10-21
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 21784
; TYPE: DNA
; ORGANISM: HUMAN
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(21784)
; OTHER INFORMATION: n = A,T,C or G
US-10-274-031-3

```

Query Match	56.7%;	Score 17;	DB 6;	Length 21784;
Best Local Similarity	100.0%;	Pred. No. 1.2e+02;		
Matches 17;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	1	CATGCCATGGGTAGAAC	17
Db	21733	CATGCCATGGGTAGAAC	21717

```

RESULT 5
US-09-724-676-27540/c
: Sequence 27540, Application US/09724676
: GENERAL INFORMATION:
: APPLICANT: Compugen LTD
: TITLE OF INVENTION: Variants of alternative splicing
: FILE REFERENCE: 129181.4 Compugen
: CURRENT APPLICATION NUMBER: US/09/724,676
: CURRENT FILING DATE: 2000-11-28
: NUMBER OF SEQ. ID NOS. 97222
: SOFTWARE: PatentIn version 3.2
: SEQ ID NO 27540
: LENGTH: 750
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-724-676-27540

```

Query Match	55.3%	Score 16.6	DB 5	Length 750
Best Local Similarity	82.6%	Pred. No. 1e+02		
Matches 19, Conservative	0	Mismatches 4	Indels 0	Gaps 0
Oy	1	CATGCCATGGGTAGACGGCGCTG	23	
db	62	CATGCCACGGGCGAGTACGGGCGAG	40	

RESULT 6
 US-10-264-237-1263
 : Sequence 1263, Application US/10264237
 : GENERAL INFORMATION:
 : APPLICANT: Birse et al.
 : TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 : FILE REFERENCE: PAl31p1
 : CURRENT APPLICATION NUMBER: US/10/264,237


```

RESULT 9
US-10-152-319A-1686
/ Sequence 1686, Application US/10152319A
/ GENERAL INFORMATION:
/ APPLICANT: Mendrick, Donna
/ APPLICANT: Porter, Mark
/ APPLICANT: Johnson, Kory
/ APPLICANT: Higgs, Brandon
/ APPLICANT: Castle, Arthur
/ APPLICANT: Elashoff, Michael
/ TITLE OF INVENTION: Molecular Toxicology Modeling
/ FILE REFERENCE: 44921-5089-US
/ CURRENT APPLICATION NUMBER: US/10/152,319A
/ PRIORITY FILING DATE: 2002-05-22
/ PRIOR APPLICATION NUMBER: US 60/292,335
/ PRIOR FILING DATE: 2001-05-22
/ PRIOR APPLICATION NUMBER: US 60/297,523
/ PRIOR FILING DATE: 2001-06-13
/ PRIOR APPLICATION NUMBER: US 60/298,925
/ PRIOR FILING DATE: 2001-06-19
/ PRIOR APPLICATION NUMBER: US 60/303,810
/ PRIOR FILING DATE: 2001-07-10
/ PRIOR APPLICATION NUMBER: US 60/303,807
/ PRIOR FILING DATE: 2001-07-10
/ PRIOR APPLICATION NUMBER: US 60/303,808
/ PRIOR FILING DATE: 2001-07-10
/ PRIOR APPLICATION NUMBER: US 60/315,047
/ PRIOR FILING DATE: 2001-08-28
/ PRIOR APPLICATION NUMBER: US 60/324,928
/ PRIOR FILING DATE: 2001-09-27
/ PRIOR APPLICATION NUMBER: US 60/330,867
/ PRIOR FILING DATE: 2001-11-01
/ PRIOR APPLICATION NUMBER: US 60/330,462

```

```
; PRIOR FILING DATE: 2001-10-22
; Remaining Prior Application data removed - See File Wrapper or PAM.
; NUMBER OF SEQ ID NOS: 2221
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1686
; LENGTH: 2626
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. NM_017245
US-10-152-319A-1686

Query Match          54.7%; Score 16.4; DB 6; Length 2626;
Best Local Similarity 76.9%; Pred. No. 1.6e+02;
Matches 20; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 GCCATGGGTAGACGGGCGTATACC 29
Db 1193 GCCATGGGTATTAGAGCTGCGACCC 1218

RESULT 10
US-10-264-237-2850/c
; Sequence 2850, Application US/10264237
; GENERAL INFORMATION:
; APPLICANT: Birste et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA131P1
; CURRENT APPLICATION NUMBER: US/10/264,237
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/16450
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/205,515
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 2876
; SOFTWARE: Patentin Ver. 3.1
; SEQ ID NO 2850
; LENGTH: 8924
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-264-237-2850

Query Match          54.7%; Score 16.4; DB 6; Length 8924;
Best Local Similarity 76.9%; Pred. No. 2e+02;
Matches 20; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CATGCCATGGGTAGACGGGCGTGATA 26
Db 5814 CATGCCAAGGAGGAGACGCTGTATA 5789

RESULT 11
US-10-264-237-2849/c
; Sequence 2849, Application US/10264237
; GENERAL INFORMATION:
; APPLICANT: Birste et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA131P1
; CURRENT APPLICATION NUMBER: US/10/264,237
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/16450
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/205,515
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 2876
; SOFTWARE: Patentin Ver. 3.1
; SEQ ID NO 2849
; LENGTH: 8925
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-264-237-2849

Query Match          54.7%; Score 16.4; DB 6; Length 8925;
```

```
Best Local Similarity 76.9%; Pred. No. 2e+02;
Matches 20; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CATGCCATGGGTAGACGGGCGTGATA 26
Db 5813 CATGCCAAGGAGGAGACGCTGTATA 5788

RESULT 12
US-10-264-213-27/c
; Sequence 27, Application US/10264213
; GENERAL INFORMATION:
; APPLICANT: Glenn, Matthew
; APPLICANT: Havukkala, Ilkka J
; APPLICANT: Lubbers, Mark William
; APPLICANT: Dekker, James
; TITLE OF INVENTION: Polynucleotides, materials incorporating
; FILE REFERENCE: 11000.1043c3
; CURRENT APPLICATION NUMBER: US/10/264,213
; CURRENT FILING DATE: 2002-10-03
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 13825
; TYPE: DNA
; ORGANISM: Lactobacillus rhamnosus
US-10-264-213-27

Query Match          54.7%; Score 16.4; DB 6; Length 13825;
Best Local Similarity 76.9%; Pred. No. 2.1e+02;
Matches 20; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 ATGCCATGGGTAGACGGGCGTGATAC 27
Db 13087 ATCCCTTGGGTAGATGGGCCGATCC 13062

RESULT 13
US-09-513-999C-1419/c
; Sequence 1419, Application US/09513999C
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclet, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: 59,US2,REG.
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 14199
; LENGTH: 192
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-513-999C-14199

Query Match          54.0%; Score 16.2; DB 5; Length 192;
Best Local Similarity 72.4%; Pred. No. 1.3e+02;
Matches 21; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 2 ATGCCATGGGTAGACGGGCGTGATACCA 30
Db 143 ATGCCATGGCAACACCGGCTGCTCCGA 115

RESULT 14
US-09-724-676-29903/c
; Sequence 29903, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
```

```
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO: 29903
; LENGTH: 1110
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (128)..(128)
; OTHER INFORMATION: n is a,c,g, or t
US-09-724-676-29903

Query Match          54.0%; Score 16.2; DB 5; Length 1110;
Best Local Similarity 72.4%; Pred. No. 1.7e+02;
Matches 21; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
```

```
QY      2 ATGCCATGGGTGAGACGGGCTGATACCCA 30
        ||||| ||||| ||||| ||||| |||||
Db      688 ATCCCAACAGTAGTGTGGGGCAGATACCCA 660
```

```
RESULT 15
US-09-724-676-29900/C
; Sequence 29900, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO: 29900
; LENGTH: 1124
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (128)..(128)
; OTHER INFORMATION: n is a,c,g, or t
US-09-724-676-29900
```

```
Query Match          54.0%; Score 16.2; DB 5; Length 1124;
Best Local Similarity 72.4%; Pred. No. 1.7e+02;
Matches 21; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
```

```
QY      2 ATGCCATGGGTGAGACGGGCTGATACCCA 30
        ||||| ||||| ||||| ||||| |||||
Db      702 ATCCCAACAGTAGTGTGGGGCAGATACCCA 674
```

Search completed: November 23, 2002, 19:40:30
Job time : 69 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 17:14:56 : Search time 2616 Seconds

(without alignments)
333.748 Million cell updates/sec

Title: US-09-848-781-4

Perfect score: 30
Sequence: 1 catgccatggtagaacggctgataccca 30

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pt:*
10: gb_ro:*
11: gb_sy:*
12: gb_un:*
13: gb_vl:*
14: em_ba:*
15: em_fun:*
16: em_hum:*
17: em_in:*
18: em_mu:*
19: em_om:*
20: em_or:*
21: em_ov:*
22: em_ph:*
23: em_pl:*
24: em_ro:*
25: em_sts:*
26: em_un:*
27: em_vl:*
28: em_htg_hum:*
29: em_htg_inv:*
30: em_htg_mus:*
31: em_htg_other:*
32: em_htg_mus:*
33: em_htg_pln:*
34: em_htg_rod:*
35: em_htg_mam:*
36: em_htg_vrt:*
37: em_hngo_hum:*
38: em_hngo_mus:*
39: em_hngo_mus:*
40: em_hngo_mus:*
41: em_hngo_mus:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	30	100.0	30	6	AX391153 Sequence
2	21	70.0	1533	6	AX413753 Sequence
3	21	70.0	1533	6	AX415653 Sequence
4	21	70.0	1952	1	LMWPL
5	21	70.0	160050	1	AL591974 L. monocyt
6	20.2	67.3	95517	2	AC105784 Listeria
7	20.2	67.3	102624	2	AL161912 Rattus no
8	20.2	67.3	109240	2	AL161912 Human DNA
9	20.2	67.3	138172	2	AC118871 Rattus no
10	19.6	65.3	200064	2	AC023067 Homo sapi
11	19.4	64.7	1215	6	AC068580 Homo sapi
12	19.4	64.7	3465	1	AX414431 Sequence
13	19.4	64.7	4562	6	AX416781 L. monocytog
14	19.4	64.7	6333	3	DMWASTER
15	19.4	64.7	35372	2	AC027302 Mus muscu
16	19.4	64.7	68425	2	AC101179 Mus muscu
17	19.4	64.7	121987	2	AC130109 Rattus no
18	19.4	64.7	170119	9	AC091970 Homo sapi
19	19.4	64.7	184335	2	AC114572 Mus muscu
20	19.4	64.7	190594	10	AL606965 Mus muscu
21	19.4	64.7	251199	9	HUAE000659 Homo sapi
22	19.2	64.0	3957	8	SCYGL207W S. cerevisia
23	19.2	64.0	4502	8	YSCDC68 M73533 yeast CDC68
24	19.2	64.0	166436	9	AL159156 Human DNA
25	19.2	64.0	227117	2	AC122341 Mus muscu
26	19.2	64.0	339044	2	AC122275 Mus muscu
27	19.2	63.3	7560	14	HPV49 X74480 Human papill
28	19.2	63.3	65698	2	AC122745 Mus muscu
29	19.2	63.3	120029	2	HSJ282810 Homo sapi
30	19.2	63.3	133217	2	AC111794 Rattus no
31	19.2	63.3	150903	2	AC113218 Rattus no
32	19.2	63.3	152966	9	AL359713 Human DNA
33	19.2	63.3	182792	2	AC098154 Rattus no
34	19.2	63.3	198820	2	AC130973 Rattus no
35	19.2	63.3	249692	2	AC125986 Rattus no
36	18.8	62.7	121217	2	AC120969 Rattus no
37	18.8	62.7	13250	9	AF513858 Homo sapi
38	18.8	62.7	20397	1	AE008870 Salmonell
39	18.8	62.7	32761	2	AC020312 Drosophil
40	18.8	62.7	44160	9	AC000049 Homo sapi
41	18.8	62.7	71141	2	AC106159 Rattus no
42	18.8	62.7	101155	9	AL449423 Human DNA
43	18.8	62.7	106295	2	AC122671 Rattus no
44	18.8	62.7	12912	2	AC123509 Rattus no
45	18.8	62.7	144092	5	AC092081 Gallus ga

ALIGNMENTS

RESULT 1
AX391153
LOCUS AX391153 30 bp DNA linear PAT 19-MAR-2002
DEFINITION Sequence 4 from Patent WO0210433.
ACCESSION AX391153
VERSION AX391153.1 GI:19584242
KEYWORDS
SOURCE
ORGANISM
Listeria monocytogenes.
Listeria monocytogenes
Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
REFERENCE
SANDERS, M.C.
METHOD AND DEVICE FOR DETECTING BACTERIAL CONTAMINATION
PATENT: WO 0210433-A 4 07-FEB-2002;
Expressive Constructs, Inc. (US)

FEATURES	
source	Location/Qualifiers 1.30
	/organism="Listeria monocytogenes" /db_xref="taxon:1639"
BASE COUNT	8 a 8 c 9 g 5 t
ORIGIN	

	Query Match	100.0%	Score 30;	DB 6;	Length 30;
	Best Local Similarity	100.0%;	Pred. No. 0.026;		
Matches	30; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0
Oy	1 CATGCCATGGGTAGAACGGGCTGATACCCA	30			
Db	1 CATGCCATGGGTAGAACGGGCTGATACCCA	30			

RESULT 2				
AX413753				
LOCUS	AX413753	1533 bp	DNA	linear
DEFINITION	Sequence 744 from Patent WO0228891.			
ACCESSION	AX413753			
VERSION	AX413753.1	GI:21446210		
KEYWORDS	.			
SOURCE	Listeria monocytogenes EGD-e.			
ORGANISM	Listeria monocytogenes EGD-e			
	Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.			
REFERENCE	1			
AUTHORS	Glaser,P. and Kunst,F.			
TITLE	Listeria innocua genome and applications			
JOURNAL	Patent: WO 0228891-A 744 11-APR-2002;			

Query Match	70.0%	Score 21	DB 6	Length 1533
Best Local Similarity	100.0%	Pred. No.	1.2e+02	
Matches 21	Conservative 0	Mismatches 0	Indels 0	Gaps 0

RESULT 3				
AX415653				
LOCUS	AX415653	1533 bp	DNA	linear
DEFINITION	Sequence 2644 from Patent WO0228891.			PAT 14-JUN-2002
ACCESSION	AX415653			
VERSION	AX415653.1	GI:21448110		
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS	Glaser, P. and Kunst, F.			
TITLE	Listeria innocua genome and applications			
JOURNAL	Patent: WO 0228891-A 2644 11-APR-2002;			
	Pasteur Institut (FR)			
FEATURES				
source	Location/Qualifiers			
	1..1533			

Query Match	70.0%	Score 21:	DB 6:	Length 1533;
Best Local Similarity	100.0%	Pred. NO.	1.2e+02;	
Matches 21:	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY 10 GGTAGACGGGCTGATACCA 30
 |||||
Db 600 GGTAACAAGGCTGATACCA 620

LOCUS	LMPL	1952 bp	DNA	linear	BCT 15-APR-1992
DEFINITION	LMPL X5.6 monocytogenes mpl gene for metalloprotease.				
ACCESSION	X54619				
VERSION	X54619.1 GI:44114				
KEYWORDS	metalloprotease; mpl gene.				
SOURCE	Listeria monocytogenes.				
ORGANISM	Listeria monocytogenes.				
REFERENCE	Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.				
AUTHORS	1 (bases 1 to 1952)				
TITLE	Domann, E.				
JOURNAL	Direct Submission				
REFERENCE	Submitted (30-OCT-1990) E. Domann, INSTITUT FUER GENETIK UND				
AUTHORS	MIKROBIOLOGIE, KOENTGENRING 11, D 8700 WUERZBURG				
TITLE	See also X15127				
JOURNAL	2 (bases 1 to 1952)				
REFERENCE	Domann, E., Leimeister-Wachter, M., Goebel, W. and Chakraborty, T.				
AUTHORS	Molecular cloning, sequencing, and identification of a				
TITLE	metalloprotease gene from Listeria monocytogenes that is species				
JOURNAL	specific and physically linked to the listeriolysin gene				
MEDLINE	Infect. Immun. 59 (1), 65-72 (1991)				
PUBMED	91100010				
COMMENT	1898903				
FEATURES	Data kindly reviewed (18-MAR-1991) by Domann E.				
SOURCE	Location/Qualifiers				
	1..1952				

```

mat_peptide      277. -1806
                 /gene="mpi"
                 /product="metalloprotease"
                 /evidence=experimental
terminator       1842. -1904
                 /gene="mpi"

BASE COUNT      687 a      390 g      559 t
ORIGIN

```

Query Match 70.0%; Score 21; DB 1; Length 1952;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 GGTAGACGGCTGATACCA 30
|||||
Db 876 GGTAGACGGCTGATACCA 896

RESULT 5
AL591974
LOCUS
DEFINITION Listeria monocytogenes strain EGD, complete genome, segment 2/12.
ACCESSION AL591974 AL591874
VERSION AL591974.1 GI:16409560
KEYWORDS
SOURCE Listeria monocytogenes.
ORGANISM Listeria monocytogenes
REFERENCE 1
AUTHORS Glaser, P., Frangeul, L., Buchrieser, C., Rusniok, C., Amend, A.,
Baquero, F., Berche, P., Bloeker, H., Brandt, P., Chakraborty, T.,
Charbit, A., Chetoui, F., Couve, E., de Daruvar, A., Dehoux, P.,
Domann, E., Dominguez-Bernal, G., Duchaud, E., Durant, L.,
Dussurget, O., Entian, K. D., Fsihi, H., Portillo, F. G., Garrido, P.,
Gautier, L., Goebel, W., Gomez-Lopez, N., Hain, T., Hauf, J.,
Jackson, D., Jones, L. M., Kaerst, U., Kreft, J., Kuhn, M., Kunst, F.,
Kurapat, G., Madueno, E., Maitournan, A., Vicente, J. M., Ng, E.,
Nedjati, H., Nordstieck, G., Novella, S., de Pablo, B., Perez-Diaz, J. C.,
Purcell, R., Remmel, B., Rose, M., Schueter, T., Simoes, N.,
Tierrez, A., Vazquez-Boland, J. A., Voss, H., Wehlund, J. and Cossart, P.
Comparative genomics of Listeria species
Science 294 (5543), 849-852 (2001)

TITLE
JOURNAL MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT
FEATURES
SOURCE

1 (bases 1 to 160050)
2 (bases 1 to 160050)
Glaser, P., Frangeul, L. and Rusniok, C.
Direct Submission
Submitted (06-JUN-2001) Glaser P., Institut Pasteur, Genomique des
Microorganismes Pathogenes, 25 rue du Docteur Roux, 75724 Paris
Cetec 15, FRANCE
E-mail: pglaser@pasteur.fr
Phone: +33 1 45 68 89 96, Fax: +33 (0)1 45 68 87 86.

Location/Qualifiers
1. 160050
/organism="Listeria monocytogenes"
/strain="EGD-e"
/db_xref="taxon:1639"
80. 402
/gene="lmo0196"
80. 85
/gene="lmo0196"
94. 402
/gene="lmo0196"
/note="similar to B. subtilis SpovG protein"
/codon_start=1
/transl_table=1
/protein_id="CAD00723.1"
/db_xref="GI:16409561"
/db_xref="SPTRMBL:O8YAD5"
/translation="METTDVRLRVEFTDGRMKAISSITIDGFEVHIDIRVINDGEGLF
VAMPSKRTPPGEFNDIAHPINSTRAKIQEVLALAYEVADEPAPNNESSADESIEVEN
"

gene 508. 876
/gene="lmo0197"
RBS 508. 513
/gene="lmo0197"
CDS 522. 830
/gene="lmo0197"
/note="similar to B. subtilis SpovG protein"
/codon_start=1
/transl_table=1
/protein_id="CAD00724.1"

terminator
gene
RBS
CDS

845. 876
/gene="lmo0197"
1202. 2590
/gene="gcaad"
1202. 1207
/gene="gcaad"
1217. 2590
/gene="gcaad"
/note="highly similar to UDP-N-acetylglucosamine
pyrophosphorylase"
/codon_start=1
/transl_table=1
/protein_id="CAD00725.1"
/db_xref="GI:16409563"
/db_xref="SPTRMBL:O8YAD4"
/translation="MSKRYAVVLAAGGCTRRKSKLYVLPYVCKPVENHYVDISTL
NVDAVVITVIGGAEKVOEHLGKSEFYKODEQLTAHAVALQAKELGKGVTLVCG
DTPLEASTMEALTKYHHEKRAKATILTYIEDPTGYGRIIRDLGIVERKIVERKAT
EKEORISEINTGYTCFENKALFEALVNSNDNVGEEYLLPDVILKIDSEVVAAYM
ESFESIGVNDRIALAEASKLMORRINENHRNGVTYVNPSTYIDIDVRIGOPTVE
PGVWLARKTYIGDCVYTSSEIYNSYIGEVHYRTSSIESPKGDDVOGTPYHLAP
ESDIDHVKITGNVYETKRAYVGEETKLPHFTYMGDAELGKRVNGCSISAVNDGKAK
AKTIGDNVVGCSNLIAPKVDRAFIAGSTITRDVDPDALGIARAKDNKLTGYA
KLNHRSK"
join(2629. 2634, 2641. 3597)
/gene="prs"
2629. 2634
/gene="prs"
2641. 3597
/gene="prs"
/codon_start=1
/transl_table=1
/product="phosphoribosyl pyrophosphate synthetase"
/protein_id="CAD00726.1"
/db_xref="GI:16409564"
/db_xref="SPTRMBL:Q48793"
/translation="MSNEYEPDKLIFSLNSNRLEAEIKAEVIGELKSSVTHFSOG
EIOINIESIRGCHVYVIOSTNSPNNQNLMEILLIMDALRKASATINIVAPYGVAR
ODKRAKRSREPTIAKLVANILTACATMTIMDMARPOIGEPDIPIDILNARLSDY
FSEHRLDDLVVSPDHGVTYRARKMADRLKAPATIDKRRPRNVAEVAIVGVNG
KVCILIDILDTAGTITLAAKLEAGATKYACCSHPVLSGPMKRIESSPLEKLV
TNSIALPEERKIDKMEQLSVAILGEAIVRVENAVSSILE"
complement(3596. 3631)
/gene="prfa"
complement(3596. 3631)
/gene="prfa"
3607. 3642
/gene="prs"
3607. 3642
/gene="prs"
complement(join(3640. 4353, 4361. 4365))
/gene="prfa"
complement(3640. 4353)
/gene="prfa"
/transl_table=1
/product="listeriolysin positive regulatory protein"
/protein_id="CAD00727.1"
/db_xref="GI:16409565"
/db_xref="SWISS-PROT:P22262"
/translation="MNAQAEFEKFLYETNGIKRPFQFKKELIFNOMDPQRCYFLVYG
ITKLTISENGTINMLQYVGAFYMGFIDLETSTGVYVNEVISEQATAYVIRINPL
KEILSKNLTREYVYFQRLQKQVYSIAKFNDSINGTSGICGLITLIVYKCEPD
GKITLDNLTMOELGSSGTAHSSAVSRIISKLEQEVYIYKKSCTFVQNDLTKRA
PKIDENFYLAACPATWGRKN"
complement(4361. 4365)
/gene="prfa"
complement(4624. 5591)

CDS /gene="p1ca"
complement(4624..5577)
/gene="p1ca"
/codon_start=1
/transl_table=11
/product="phosphatidylinositol-specific phospholipase c"
/protein_id="CAD00728.1"
/db_xref="GI:16409566"
/db_xref="SWISS-PROT:P34024"
/translation="MYKNLQRLVLLLCFLLYFPFPLGKAYSLNNMKPKIKSVT
TKOMKSLPDTNIALSTPGHDMSTYGDITWTLKRLATQRMISLQCEAGIRY
IDIRAKDNITHTGFLFTLASLGVLETTOTFLKPKPKETITMRKDEONSDSDYR
IQPLNITYKDYFTPTPTDSNKLPTLKQVRKILLLSNHTRKPLVINSRKFGQGF
ANOVYODYNGPSVTKFKFEIVQTAQSKADNKLFLNHSATSLTFTPROYAAALN
NKVEQFVLNTESEKVGILIMDFPEKQTIKNIKNNFN"
complement(5586..5591)
/gene="p1ca"
RBS 5804..7408
/gene="hly"
RBS 5804..5809
/gene="hly"
CDS 5819..7408
/gene="hly"
/codon_start=1
/transl_table=11
/product="listeriolysin O precursor"
/protein_id="CAD00729.1"
/db_xref="GI:16409567"
/db_xref="SWISS-PROT:P13128"
/translation="MKIMLVFTITLVLSPLAQOTEAKDASAFNKENSISMAFPAS
PASRPKPLEKHADEIDISLYGLDYNKNNVLYHGDAVTNPVPRKQKDGMEYIVE
KKKKSINONNADIOVYNAISLTYPGALVKANSELVENDPVLVPRKSLTSLDLP
MTNQDNKIYVKNATKSNMNAVNTLVERNEKYAOAYPAVSAKIDYDDMAKSESLI
AFEGFAKAVNNSLNVNFGAISEGKMOEVEISFKOYYVNVNNEPDRPSPKAVTK
EOLQAGVNAENPVAISSVAYGRQVYLKLSNHSSTKYKAFDAVASGKSVSGVEL
TNIIKNSFKAIYGSADVOIIDNIGDLRDLKKAATNRETPGPIATYTNFL
KNELAVIKNNSEITETTSKATYDGIKINIDHSGGVVAQFNISWDEYNIDPENEIYOH
KMWSENNRSKLHFTSSITLPGNARINIVAYACEGCLAMEWMRTVIDDNLPLVNRN
ISINGTLYTPKRSNKVNDPNE"
7727..9271
/gene="mpl"
RBS 7727..7732
/gene="mpl"
CDS 7739..9271
/gene="mpl"
/codon_start=1
/transl_table=11
/product="zinc metalloproteinase precursor"
/protein_id="CAD00730.1"
/db_xref="GI:16409568"
/db_xref="SWISS-PROT:P23224"
/translation="MKSLLICIIIVIAFOAHFTMTVKADSVGEKLNONTOKKTPAD
LKALPDCARKEFKNFRLDKDKICVTHYTLALSSGYLTDDEIKVHTPPNKI
TFINDLQOGLRTINQIKITEKNAIEKFAIEGSEARHSTYVGNPKETIILNSR
TKRLVYNIKLFAEPVAVSVIVOVDFGAILKKONMLSEVRADTHDFOALGGAN
RLQPLVHMKINDFLVYDRTNKGILRTFDLKHNTDTSFGKVSNKNNMFTDPFSS
ADAHFYASEVYEVYKVVLESIDSGEGEISFYHGLNCNNAMWDEGIEIYGGDK
KNKPEPSGAKTIVGHELTFAVIOYASGLEEGOSGALNNSFADVGFAPLPHWILGE
DYCVGSRGDRIRSTIKDPKYNQAAHMDYESTLPTREEDMGGVYNSGIPKAYNT
ITKLGKERTQQLYFRALKYITLKKSQFTDAKALQAAKIDGEDASKKVAWEAVG
VN"
9456..11389
/gene="acta"
RBS 9456..9461
/gene="acta"

Query Match 70.0%; Score 21; DB 1; Length 160050;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 GGTAGAACGGGCTGATACCA 30
|||||
DB 8338 GGTAGAACGGGCTGATACCA 8358

QY 10 GGTAGAACGGGCTGATACCA 30
|||||
DB 8338 GGTAGAACGGGCTGATACCA 8358

RESULT 6
AC105784
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE

AUTHORS

1 (bases 1 to 95517)

Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,

Alsbrooks,S.L., Amaralunge,H.C., Are,J.R., Ayale,M., Banks,T.,

Barbata,J., Benton,J., Bimaye,K., Blankenburg,K., Bonin,D.,

Bouck,J., Bowie,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P.,

Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,

Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,

Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,

Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,

Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,

Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinu,H.H.,

Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,

Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,

Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,

Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,

Gorell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,

Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,

Hernandez,O., Hodgson,A., Hogue,M., Holloway,C., Hollins,B.,

Homs,F., Howard,S., Huber,J., Hulik,S., Hume,J., Jackson,L.E.,

Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,

Karlsson,E., Kelly,S., Khan,U., King,L., Korvan,J., Kovar,C.,

Kratovic,J., Kurshl,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,

Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Louisedge,H.,

Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,

Maheshwari,M., Mapa,P., Martin,R., Martindale,A., Martinez,E.,

Masey,E., Mawhney,E., McLeod,M.P., Meador,M., Mel,G., Metzger,M.,

Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,

Moser,M., Neal,D., Newton,N., Newton,N., Nguyen,N., Nguyen,N.,

Nguyen,N., Nickerson,E., Nwokenkwu,S., Ogih,M., Okwuonu,G.,

Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,

Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,

Rives,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savery,G.,

Scherer,S., Scott,G., Shen,H., Shooshari,N., Sisson,I.,

Sodergren,E., Sonaite,T., Sparks,A., Stanley,H., Stone,H.,

Sutton,A., Svatek,A., Taber,P., Tamerisa,A., Tamerisa,K., Tang,H.,

Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,

Umsni,K., Vasquez,L., Vera,Y., Villalon,D., Vinson,R., Wang,O.,

Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,

Williams,G., Williamson,A., Wleczek,R., Woodson,S., Worley,K.,

Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorilla,S., Nelson,D.,

Wu,stock,G. and Gibbs,R.

Unpublished

2 (bases 1 to 95517)

Worley,K.C.

Direct Submission

Submitted (10-JAN-2002) Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 95517)

Worley,K.C.

Direct Submission

Submitted (13-JUL-2002) Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

On Jul 12, 2002 this sequence version replaced gi:18105075.

----- Genome Center

Center: Baylor College of Medicine

```

Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information -----
Center project name: GNRE
Center clone name: CH230-249L22
----- Summary Statistics -----
Sequencing vector: Plasmid:
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 64678 bases at least Q40
Consensus quality: 67380 bases at least Q30
Consensus quality: 69868 bases at least Q20
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/gendank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 42 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
1178: contig of 1178 bp in length
1179 1278: gap of unknown length
1279 2388: contig of 1110 bp in length
2389 2488: gap of unknown length
2489 3622: contig of 1134 bp in length
3623 3722: gap of unknown length
3723 4925: contig of 1203 bp in length
4926 5025: gap of unknown length
5026 6559: contig of 1534 bp in length
6560 6659: gap of unknown length
6660 8066: contig of 1407 bp in length
8067 8166: gap of unknown length
8167 9580: contig of 1414 bp in length
9581 9680: gap of unknown length
9681 10751: contig of 1071 bp in length
10752 10851: gap of unknown length
10852 12287: contig of 1436 bp in length
12288 12387: gap of unknown length
12388 13787: contig of 1400 bp in length
13788 13887: gap of unknown length
13888 15284: contig of 1397 bp in length
15285 15384: gap of unknown length
15385 16829: contig of 1445 bp in length
16830 16929: gap of unknown length
16930 18932: contig of 2003 bp in length
18933 19032: gap of unknown length
19033 20837: contig of 1805 bp in length
20838 20937: gap of unknown length
20938 22065: contig of 1128 bp in length
22066 22165: gap of unknown length
22166 23511: contig of 1346 bp in length
23512 23611: gap of unknown length
23612 25864: contig of 2253 bp in length
25865 25964: gap of unknown length
25965 27439: contig of 1475 bp in length
27440 27539: gap of unknown length
27540 28990: contig of 1451 bp in length
28991 29090: gap of unknown length
29091 30987: contig of 1897 bp in length
30988 31087: gap of unknown length
31088 32838: contig of 1751 bp in length
32839 32938: gap of unknown length
32939 35167: contig of 2229 bp in length
35168 35267: gap of unknown length
35268 37456: contig of 2189 bp in length
37457 37556: gap of unknown length
37557 39990: contig of 2434 bp in length
39991 40090: gap of unknown length
40091 41497: contig of 1407 bp in length

```

```

* 41498 41597: gap of unknown length
* 41598 44130: contig of 2533 bp in length
* 44131 44230: gap of unknown length
* 44231 46359: contig of 2129 bp in length
* 46360 46459: gap of unknown length
* 46380 48380: contig of 1921 bp in length
* 48381 48480: gap of unknown length
* 48481 51052: contig of 2572 bp in length
* 51053 51152: gap of unknown length
* 51153 53362: contig of 2210 bp in length
* 53363 53462: gap of unknown length
* 53463 55699: contig of 2237 bp in length
* 55700 55799: gap of unknown length
* 55800 58296: contig of 2497 bp in length
* 58297 58396: gap of unknown length
* 58397 60734: contig of 2338 bp in length
* 60735 60834: gap of unknown length
* 60835 63175: contig of 2341 bp in length
* 63176 63275: gap of unknown length
* 63276 66079: contig of 2804 bp in length
* 66080 66179: gap of unknown length
* 66180 68066: contig of 1887 bp in length
* 68067 68166: gap of unknown length
* 68167 71491: contig of 3325 bp in length
* 71492 71591: gap of unknown length
* 71592 76081: contig of 4490 bp in length
* 76082 76181: gap of unknown length
* 76182 79677: contig of 3496 bp in length
* 79678 79777: gap of unknown length
* 79778 84866: contig of 5089 bp in length
* 84867 84966: gap of unknown length
* 84967 88530: contig of 3564 bp in length
* 88531 88630: gap of unknown length
* 88631 95517: contig of 6887 bp in length.
* Location/Qualifiers
* source
* 1..95517
* /organism="Rattus norvegicus"
* /db_xref="taxon:10116"
* /clone="CH230-249L22"
BASE COUNT 21181 a 24491 c 23626 g 21010 t 5209 others
ORIGIN
Query Match 67.3% Score 20.2; DB 2; Length 95517;
Best Local Similarity 88.0%; Pred. No. 1.3e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 5 CCATGGGTAGAACGGCGTATACCC 29
||||| ||||| ||||| |||
Db 3320 CCATGGGAGAACGTCGTATGCC 3344
RESULT 7
AL161912 102624 bp DNA linear PRI 04-DEC-2001
LOCUS Human DNA sequence from clone Rpl1-403N16 on chromosome 9 Contains
DEFINITION a novel pseudogene, complete sequence.
ACCESSION AL161912
VERSION AL161912.15 GI:14270137
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 102624)
AUTHORS Mashreghi-Mohammadi, M.
TITLE Direct Submission
JOURNAL Submitted (04-DEC-2001) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquerry@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk
On May 31, 2001 this sequence version replaced gi:14132858.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the

```

variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at

http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping Group. Further information can be found at

http://www.sanger.ac.uk/HGP/Chr9

Rp11-403N16 is from the library RPC1-11.2 constructed by the group of Pieter de Jong. For further details see

http://www.chori.org/bacpac/home.htm

VECTOR: pBACe3.6

IMPORTANT: This sequence is not the entire insert of clone Rp11-403N16. It may be shorter because we sequence overlapping sections only once, except for a short overlap.

The true right end of clone Rp11-403N16 is at 102624 in this sequence. The true left end of clone Rp11-327L9 is at 48008 in this sequence. The true right end of clone Rp11-79D8 is at 100 in this sequence.

FEATURES

```
source
1..102624
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="9"
/clone="Rp11-403N16"
/clone_lib="RPC1-11.2"
3..399
/note="MST-INTERNAL repeat: matches 530..1007 of consensus"
400..720
/note="AluY repeat: matches 11..306 of consensus"
721..897
/note="MST-INTERNAL repeat: matches 357..530 of consensus"
1245..1563
/note="MSTB-internal repeat: matches 1..326 of consensus"
1564..2003
/note="MSTB repeat: matches 4..426 of consensus"
complement(2186..2569)
/note="match: STS: Em:HS303XA9"
2290..2337
/note="MUT2E repeat: matches 348..394 of consensus"
2339..2378
/note="20 copies 2 mer tg 97% conserved"
2380..2725
/note="MUT2E repeat: matches 1..357 of consensus"
2757..2861
/note="L2 repeat: matches 2323..2430 of consensus"
2862..2895
/note="17 copies 2 mer aa 91% conserved"
3406..3553
/note="L2 repeat: matches 1590..1731 of consensus"
3876..4065
/note="L1PA16 repeat: matches 5968..6157 of consensus"
4085..9542
/note="L1PA15 repeat: matches 647..6157 of consensus"
9639..9950
/note="L1PA13 repeat: matches -178..156 of consensus"
10062..10439
/note="L1PA13 repeat: matches -667..-305 of consensus"
10902..11028
/note="FLAM_C repeat: matches 1..127 of consensus"
11244..11521
repeat_region
```

```
/note="AluX repeat: matches 1..280 of consensus"
11522..11633
/note="TIGCR1 repeat: matches 2069..2182 of consensus"
11643..12982
/note="TIGCR1 repeat: matches 244..1625 of consensus"
12983..13022
/note="HSMAR1 repeat: matches 2..41 of consensus"
13020..13135
/note="MER2 repeat: matches 1..113 of consensus"
13169..13426
/note="TIGCR1 repeat: matches 1..246 of consensus"
13427..13601
/note="L1M3 repeat: matches 5613..5787 of consensus"
13596..13880
/note="L1 repeat: matches 3577..3865 of consensus"
13879..14475
/note="L1M4C repeat: matches 1431..2030 of consensus"
14470..14657
/note="L1M4C repeat: matches 952..1139 of consensus"
15269..15304
/note="L1M3d repeat: matches 251..286 of consensus"
18538..19794
/note="L1PA3 repeat: matches 2161..3429 of consensus"
19794..22500
/note="L1PA3 repeat: matches 3439..6146 of consensus"
22762..22830
/note="Charlie3 repeat: matches 1..75 of consensus"
22831..23621
/note="L1M4 repeat: matches 3922..4689 of consensus"
24042..24666
/note="L2 repeat: matches 2072..2687 of consensus"
24741..24886
/note="MER58C repeat: matches 36..59 of consensus"
26049..26133
/note="L1PA15-16 repeat: matches 797..891 of consensus"
26134..26426
/note="AluS9 repeat: matches 16..303 of consensus"
26427..26774
/note="L1PA15-16 repeat: matches 891..1176 of consensus"
26716..26820
/note="match: GSS: Em:AQ180515"
26729..26856
/note="match: GSS: Em:AQ52470"
26731..26820
/note="match: GSS: Em:AQ799534"
26732..26924
/note="match: GSS: Em:AQ710665"
26757..26857
/note="match: GSS: Em:AQ053667"
26768..26908
/note="match: GSS: Em:AQ817478"
26774..26856
/note="match: GSS: Em:AQ493641"
28139..28262
/note="L1PA6 repeat: matches 6020..6143 of consensus"
28263..29347
/note="MER11C repeat: matches 1..1071 of consensus"
29348..29451
/note="L1PA7 repeat: matches 5939..6029 of consensus"
29886..29962
/note="WIR repeat: matches 21..99 of consensus"
30149..33769
/gene="BA403N16.1"
/pseudo
30149..33769
/gene="BA403N16.1"
/note="BA403N16.1 (novel protein)
match: proteins: Tr:Q9NZP6 Tr:Q26054 Tr:Q18975 Tr:Q9S728
Tr:Q9I0M1 Tr:Q22807 Tr:Q9Y2A2 Tr:Q80744 Sw:Q06831
Tr:Q9XDH2 Tr:Q9Y957 Tr:Q99495 Sw:P29400 Sw:P54259
Tr:Q9KY45"
/pseudo
/codon_start=1
repeat_region
```

misc_feature	/evidence=not_experimental 31378. .31513 /gene="DA403N16.1" /note="match: STS: Em:HSJ34F11" /pseudo 31521. .31843 /gene="DA403N16.1" /note="match: STS: Em:HSJ34F11" /pseudo 33944. .34302 /note="match: GSS: Em:AQ237368" 34257. .34324 /note="34 copies 2 mer tg 97% conserved" 35388. .35626 /note="t2 repeat: matches 2453. .2692 of consensus" 36201. .36686 /note="t2 repeat: matches 2255. .2750 of consensus" 36903. .37253 /note="MLT1A2 repeat: matches 4. .358 of consensus" 37254. .37667 /note="MLT2A repeat: matches 1. .418 of consensus" /complement(37604. .37746) /note="match: STS: Em:LI1775" 37628. .37708 /note="match: GSS: Em:AZ354106" /complement(37631. .37725) /note="match: GSS: Em:AQ658518" /complement(37632. .37730) /note="match: GSS: Em:AZ041633" /complement(37634. .37746) /note="match: STS: Em:G08235" 37639. .37730 /note="match: STS: Em:G08863" /complement(37642. .37730)
--------------	--

Query Match	67.38;	Score 20.2;	DB 9;	Length 102624;
Best Local Similarity	88.08;	Pred. No. 1.3e+02;		
Matches 22; Conservative	0;	Mismatches 3;	Indels 0;	Gaps 0;

```

QY      2  ATGCCATGGGTAGAACGGGCTGATA  26
          ||||| ||| ||||| ||||| |||
Db 27458 ATGCCGTGGTTAGAAAGGGCTGATA 27482

```

RESULT 8	AC118871/c	AC118871	109240 bp	DNA	1linear	HTG 17-JUL-2007
LOCUS		Rattus norvegicus clone CH230-487N17.	***	SEQUENCING IN PROGRESS		
DEFINITION		***, 38 unordered pieces.				
ACCESSION		AC118871				
VERSION		AC118871.2	GI:21747340			
KEYWORDS		HTG; HTGS_PHASE1.				
SOURCE		Norway rat.				

REFERENCE
AUTHORS

1 (bases 1 to 109240)

Muzny, D.M., Adams, C., Adio-Oduola, B., Allisman, F.R., Allen, C., Alsthorpe, S.L., Amaratunga, H.C., Are, J.R., Ayale, M., Banks, T., Barbarks, J., Benton, J., Bimoge, K., Blankenbiller, K., Bonin, D., Bouck, J., Bowles, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C., Burck, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chowhry, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C., Davy-Carroll, L., Dedecich, D.A., Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Earhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorelli, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,

TITLE	JOURNAL
REFERENCE	Unpublished
AUTHORS	2 (bases 1 to 109240)
TITLE	Worley, K.C.
JOURNAL	Direct Submission
REFERENCE	Submitted (21-APR-2002) Human Genome Sequencing Center, Department
AUTHORS	of Molecular and Human Genetics, Baylor College of Medicine, One
TITLE	Baylor Plaza, Houston, TX 77030, USA
JOURNAL	3 (bases 1 to 109240)
REFERENCE	Worley, K.C.
AUTHORS	Direct Submission
TITLE	Submitted (17-JUL-2002) Human Genome Sequencing Center, Department
JOURNAL	of Molecular and Human Genetics, Baylor College of Medicine, One
REFERENCE	Baylor Plaza, Houston, TX 77030, USA
AUTHORS	On Jul 14, 2002 this sequence version replaced g1:20258335.
TITLE	
JOURNAL	
COMMENT	

Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hogue,M., Holloway,C., Hollins,B., Homtsi,F., Howard,S., Huber,J., Hulik,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolyvet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovac,C., Kratovic,J., Kureishi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichteige,O., Lieu,C., Liu,J., Liu,M., Louisgeed,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Masseswari,M., Mapa,P., Martin,R., Martindale,A., Martinez,E., Massy,E., Mawlinye,E., Mcloed,M.P., Meador,M., Mei,G., Metker,M., Miner,G., Mner,Z., Mitchell,T., Mohabbat,K., Morgan,A., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Ogunye,N., Nickerson,E., Nwokkenko,S., Oguh,M., Okunonu,G., Ogrunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L., Qullis,M., Ren,Y., Rivers,M., Rojas,A., RojuboKan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shoshitari,N., Sisson,I., Sodergren,E., Soneike,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansay,U., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,A., Villalón,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wleceyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D., Weinstein,G., and Gibbs,R.

Unpublished
Direct Submission
2 (bases 1 to 109240)
Worley,K.C.

Direct Submission
Submitted (21-APR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 109240)
Worley,K.C.

Direct Submission
Submitted (17-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Jul 14, 2002 this sequence version replaced gi:20258335.

----- Genome Center -----
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information -----
Center project name: GVXX
Center clone name: CH230-487N17
----- Summary Statistics -----
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 40767 bases at least Q40
Consensus quality: 45098 bases at least Q30
Consensus quality: 47548 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
* (see <http://www.hgsc.bcm.tmc.edu/docs/genbank.draft.data.html>)
* NOTE: This is a 'working draft' sequence. It currently
* consists of 58 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1
1159: contig of 1159 bp in length
1160 1259: gap of unknown length
1260 2317: contig of 1058 bp in length
2318 2417: gap of unknown length
2418 3420: contig of 1003 bp in length
3421 3520: gap of unknown length
3521 4967: contig of 1447 bp in length

4968 5067: gap of unknown length
5068 6253: contig of 1186 bp in length
6254 6353: gap of unknown length
6354 7521: contig of 1168 bp in length
7522 7621: gap of unknown length
7622 9115: contig of 1494 bp in length
9116 9215: gap of unknown length
9216 10280: contig of 1065 bp in length
10281 10380: gap of unknown length
10381 11492: contig of 1112 bp in length
11493 11592: gap of unknown length
11593 13135: contig of 1543 bp in length
13136 13235: gap of unknown length
13236 14951: contig of 1716 bp in length
14952 15051: gap of unknown length
15052 16515: contig of 1464 bp in length
16516 17621: gap of unknown length
17622 17721: contig of 1006 bp in length
17722 19112: gap of unknown length
19113 19212: gap of unknown length
19213 20781: contig of 1569 bp in length
20782 22310: gap of unknown length
22311 22310: contig of 1329 bp in length
22312 23770: gap of unknown length
23771 23870: gap of unknown length
23871 25203: contig of 1333 bp in length
25204 27023: gap of unknown length
27024 27123: contig of 1720 bp in length
27124 29001: gap of unknown length
29002 29101: gap of unknown length
29102 31009: contig of 1908 bp in length
31010 31109: gap of unknown length
31110 32657: contig of 1548 bp in length
32658 32757: gap of unknown length
32758 33891: contig of 1134 bp in length
33892 33991: gap of unknown length
33992 35567: contig of 1576 bp in length
35568 35667: gap of unknown length
35668 37348: contig of 1681 bp in length
37349 37448: gap of unknown length
37450 39052: contig of 1604 bp in length
39053 39152: gap of unknown length
39153 40615: contig of 1463 bp in length
40616 40715: gap of unknown length
40716 41807: contig of 1092 bp in length
41808 41907: gap of unknown length
41909 42967: contig of 1060 bp in length
42968 43067: gap of unknown length
43068 44362: contig of 1295 bp in length
44363 44462: gap of unknown length
44464 45787: contig of 1325 bp in length
45788 45887: gap of unknown length
45889 47612: contig of 1725 bp in length
47613 47712: gap of unknown length
47714 50040: contig of 2328 bp in length
50041 50140: gap of unknown length
50141 51965: contig of 1825 bp in length
51966 52065: gap of unknown length
52066 53879: contig of 1814 bp in length
53880 53979: gap of unknown length
53980 55226: contig of 1247 bp in length
55227 55326: gap of unknown length
55328 56664: contig of 1338 bp in length
56665 56764: gap of unknown length
56766 58507: contig of 1743 bp in length
58508 58607: gap of unknown length
58609 59975: contig of 1368 bp in length
59976 60075: gap of unknown length
60076 62390: contig of 2315 bp in length
62391 62490: gap of unknown length

62491 64033: contig of 1543 bp in length
64034 64133: gap of unknown length
64134 66230: contig of 2097 bp in length
66231 66330: gap of unknown length
66331 68776: contig of 2446 bp in length
68777 68876: gap of unknown length
68877 69887: contig of 1011 bp in length
69888 69987: gap of unknown length
69989 71704: contig of 1717 bp in length
71705 71804: gap of unknown length
71805 74871: contig of 3067 bp in length
74872 74971: gap of unknown length
74972 76269: contig of 1298 bp in length
76270 76369: gap of unknown length
76370 78742: contig of 2373 bp in length
78743 78842: gap of unknown length
78843 81306: contig of 2464 bp in length
81307 81406: gap of unknown length
81407 83441: contig of 2035 bp in length
83442 83541: gap of unknown length
83542 86403: contig of 2862 bp in length
86404 86503: gap of unknown length
86504 90310: contig of 3807 bp in length
90311 90410: gap of unknown length
90411 93365: contig of 2935 bp in length
93366 93465: gap of unknown length
93466 96148: contig of 2683 bp in length

Query Match 67.3%; Score 20.2; DB 2; Length 109240;
Best Local Similarity 88.0%; Pred. No. 1.3e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 5 CCATGGGTAGAACGGCTGATACC 29
Db 107807 CCATGGGAGAACGCTGATGCC 107783

||||| ||||| ||||| |||

RESULT 9
AC023067
LOCUS
DEFINITION Homo sapiens clone RP11-31D18, WORKING DRAFT SEQUENCE, 40 unordered pieces.
AC023067
VERSION AC023067.3 GI:7801428
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 138172)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens, clone RP11-31D18
Unpublished
2 (bases 1 to 138172)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N., Anderson,S., Baldwin,J., Barna,N., Beckert,J., Beda,F., Boguslavsky,L., Boukhalter,B., Brown,A., Burkett,G., Castle,A., Choepel,Y., Colangelo,A., Collins,S., Collymore,A., Cooke,P., Dearliano,K., Dewar,K., Domino,M., Doyle,M., Feneator,J., Ferreira,P., Fitzhugh,W., Forrest,C., Gage,D., Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heatford,A., Horton,L., Howland,J., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J., Landers,T., Lehoczy,J., Levine,R., Lieu,C., Liu,G., Locke,K., MacDonald,P., Margulis,N., McEwan,P., McGurk,A., McKernan,K., McPheeters,R., Meldrum,J., Meneus,L., Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., Olivari,T.M., Peterson,K., Pierre,N., Pisan,C., Pollara,V., Raymond,C., Riley,R., Rothman,D., Roy,A., Santos,R., Severy,P., Spencer,B., Strange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J., Tirrell,A., Vassiliev,H., Vtel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
TITLE Direct Submission
JOURNAL Submitted (08-FEB-2000) Whitehead Institute/MIT Center for Genome

COMMENT

Research, 320 Charles Street, Cambridge, MA 02141, USA
On May 14, 2000 this sequence version replaced gl:7229802.
All repeats were identified using RepeatMasker:

Smith, A.F.A. & Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: MIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

Project Information

Center project name: L4926
Center clone name: 31.D_18

Summary Statistics

Sequencing vector: M13; M7815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 120256 bases at least Q40
Consensus quality: 129429 bases at least Q30
Consensus quality: 132609 bases at least Q20
Insert size: 73000; agarose-fp
Insert size: 134272; sum-of-contigs
Quality coverage: 5.6 in Q20 bases; agarose-fp
Quality coverage: 3.0 in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently
consists of 40 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.

1 1129: contig of 1129 bp in length
* 1130 1229: gap of 100 bp
* 1230 2249: contig of 1020 bp in length
* 2250 2349: gap of 100 bp
* 2350 3469: contig of 1120 bp in length
* 3470 3569: gap of 100 bp
* 3570 5002: contig of 1433 bp in length
* 5003 5102: gap of 100 bp
* 5103 6280: contig of 1178 bp in length
* 6281 6380: gap of 100 bp
* 6381 7816: contig of 1436 bp in length
* 7817 7916: gap of 100 bp
* 7917 8991: contig of 1075 bp in length
* 8992 9091: gap of 100 bp
* 9092 10387: contig of 1296 bp in length
* 10388 10487: gap of 100 bp
* 10488 11767: contig of 1280 bp in length
* 11768 11867: gap of 100 bp
* 11868 12953: contig of 1086 bp in length
* 12954 13053: gap of 100 bp
* 13054 15567: contig of 2514 bp in length
* 15568 15667: gap of 100 bp
* 15668 17169: contig of 1502 bp in length
* 17170 17269: gap of 100 bp
* 17270 18951: contig of 1682 bp in length
* 18952 19051: gap of 100 bp
* 19052 21403: contig of 2352 bp in length
* 21404 21503: gap of 100 bp
* 21504 23637: contig of 2134 bp in length
* 23638 23737: gap of 100 bp
* 23738 25254: contig of 1517 bp in length
* 25255 25354: gap of 100 bp
* 25355 28223: contig of 2869 bp in length
* 28224 28323: gap of 100 bp
* 28324 29825: contig of 1502 bp in length
* 29826 29925: gap of 100 bp
* 29926 32235: contig of 2310 bp in length
* 32236 32335: gap of 100 bp
* 32336 34601: contig of 2266 bp in length
* 34602 34701: gap of 100 bp

FEATURES

SOURCE

* 34702 36484: contig of 1783 bp in length
* 36485 36584: gap of 100 bp
* 36585 40023: contig of 3339 bp in length
* 40024 40123: gap of 100 bp
* 40124 43905: contig of 3782 bp in length
* 43906 44005: gap of 100 bp
* 44006 46514: contig of 2609 bp in length
* 46515 46614: gap of 100 bp
* 46615 49263: contig of 2649 bp in length
* 49264 49363: gap of 100 bp
* 49364 53473: contig of 4110 bp in length
* 53474 53573: gap of 100 bp
* 53574 56741: contig of 3168 bp in length
* 56742 56841: gap of 100 bp
* 56842 61556: contig of 4715 bp in length
* 61557 61656: gap of 100 bp
* 61657 66160: contig of 4504 bp in length
* 66161 66260: gap of 100 bp
* 66261 71835: contig of 5575 bp in length
* 71836 71935: gap of 100 bp
* 71936 75219: contig of 3284 bp in length
* 75220 75319: gap of 100 bp
* 75320 80297: contig of 4578 bp in length
* 80298 80397: gap of 100 bp
* 80398 85279: contig of 4882 bp in length
* 85280 85379: gap of 100 bp
* 85380 92861: contig of 7282 bp in length
* 92862 92961: gap of 100 bp
* 92962 98940: contig of 5979 bp in length
* 98941 99040: gap of 100 bp
* 99041 103763: contig of 4723 bp in length
* 103764 103863: gap of 100 bp
* 103864 110295: contig of 6632 bp in length
* 110296 110395: gap of 100 bp
* 110396 118613: contig of 8218 bp in length
* 118614 118713: gap of 100 bp
* 118714 126224: contig of 7511 bp in length
* 126225 126324: gap of 100 bp
* 126325 138172: contig of 11848 bp in length.

LOCATION/QUALIFIERS

1. 138172
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="RP11-31D18"
/clone_lib="RPCT-11 Human Male BAC"
1. 1129
/note="assembly_fragment"
1230. 2249
/note="assembly_fragment"
2350. 3469
/note="assembly_fragment"
3570. 5002
/note="assembly_fragment"
5103. 6280
/note="assembly_fragment"
6381. 7816
/note="assembly_fragment"
7917. 8991
/note="assembly_fragment"
vector_side:left"

misc_feature
10468. 11767
/note="assembly_fragment"
11868. 12953
/note="assembly_fragment"
13054. 15567
/note="assembly_fragment"
13668. 17169
/note="assembly_fragment"
17270. 18951
/note="assembly_fragment"

misc_feature
9092. 10387
/note="assembly_fragment"
10468. 11767
/note="assembly_fragment"
11868. 12953
/note="assembly_fragment"
13054. 15567
/note="assembly_fragment"
13668. 17169
/note="assembly_fragment"
17270. 18951
/note="assembly_fragment"

misc_feature	19052. .21403
misc_feature	/note="assembly_fragment"
misc_feature	21504. .23637
misc_feature	/note="assembly_fragment"
misc_feature	23738. .25254
misc_feature	/note="assembly_fragment"
misc_feature	25353. .28223
misc_feature	/note="assembly_fragment"
misc_feature	28324. .29825
misc_feature	/note="assembly_fragment"
misc_feature	29926. .32235
misc_feature	/note="assembly_fragment"
misc_feature	32336. .34601
misc_feature	/note="assembly_fragment"
misc_feature	34702. .36484
misc_feature	/note="assembly_fragment"
misc_feature	36585. .40023
misc_feature	/note="assembly_fragment"
misc_feature	40124. .43905
misc_feature	/note="assembly_fragment"

Query Match	67.38;	Score 20.2;	DB 2;	Length 138172;
Best Local Similarity	88.0%;	Pred. No. 1.2e+02;		
Matches 22; Conservative	0;	Mismatches 3;	Indels 0;	Gaps 0;

QY 2 ATGCCATGGGTAGAACGGGCTGATA 26
|||||
Db 128491 ATGCCGTGGTTAGAAAGGGCTGATA 128515
|||||

LOCUS	AC068580	20064 bp	DNA linear	HTG 14-AUG-2007
DEFINITION	Homo sapiens chromosome 11 clone RP11-295K3 map 11, *** SEQUENCING IN PROGRESS ***. 5 unordered pieces.			

```

VERSION      AC068580.13  GI:22218576
KEYWORDS     HTG; HTGS_PHASE1; HTGS_FUZZTOP; HTGS_ACTIVEFIN.
SOURCE       human.

```

Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:
 Mammalia: Eutheria: Primates: Catarrhini: Homnidae: Homo:
 1 (bases 1 to 200064)
 Birren,B., Nusbaum,C. and Lander,E.
 Homo sapiens chromosome 11, clone RP11-295K3
 Unpublished
 2 (bases 1 to 200064)
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
 AUTHORS

TITLE Direct Submission
JOURNAL Submitted (04-MAY-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE 3 (bases 1 to 200064)
AUTHORS Birren, B., Nussbaum, C., Landø, E., Ali, A., Allen, N., Anderson, S.,

Barnes,N., Bastien,V., Bloom,T., Boguslavskiy,L., Bouhagalter,B.,
 Camaratz,J., Chang,J., Chazaroi,B., Choepel,Y., Collymore,A.,
 Cook,A., Cooke,P., DeArlellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
 Faro,S., Ferreira,P., Fittzerald,M., Gage,D., Galagan,T.,
 Gaidyna,S., Gird,S., Graham,L., Grand-Pierre,N., Hagos,B.,
 Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kemat,A.,
 Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
 Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,C.,
 McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mleenga,V.,
 Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
 O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
 Phunhkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
 Roman,J., Roy,A., Schauer,S., Schumack,R., Seaman,S., Severy,P.,
 Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
 Testfay,S., Theodore,J., Totham,K., Travers,M., Vassiliev,H.,
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
 Zemlek,L., Zimmer,A. and Zody,M.

TITLE
 JOURNAL
 COMMENT

Direct Submission
 Submitted (14-AUG-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Aug 14, 2002 this sequence version replaced gi:22123033.

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>
 Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: MIBR
 Web site: <http://www.seq.wi.mit.edu>
 Contact: sequence_submissions@genome.wi.mit.edu
 Project Information
 Center project name: LI0315
 Center clone name: 295_K_3

```

* NOTE: This is a 'working draft' sequence. It currently
* consists of 5 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of 'N', but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1
* 120033: contig of 120033 bp in length
* 120034 120133: gap of 100 bp
* 120134 135782: contig of 15649 bp in length
* 135783 135882: gap of 100 bp
* 135883 142385: contig of 6503 bp in length
* 142386 142485: gap of 100 bp
* 142486 196904: contig of 54419 bp in length
* 196905 197004: gap of 100 bp
* 197005 200066: contig of 3060 bp in length.

```

```

FEATURES
source      location/Qualifiers
1. .200064
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11"
/clone="RP11-295K3"
/clone_lib="RPCT-11 Human Male BAC"
BASE COUNT 46421 a 53023 c 50690 g 49438 t 492 others
ORIGIN

```

	65.3%	Score 19.6	DB 2:	Length 200064;
Query Match Similarity	84.6%	Pred. No. 1e+02;		
Best Local Similarity				
Matches 22; Conservative	0;	Mismatches 4;	Indels 0;	Gaps 0;
QY	5 CCATGGGTAGAACGCGCTGATACCA	30		
Db 110620 CCCTGGTGAACAGATGAACCCCA	110645			

RESULT 11	
AX414431	
LOCUS	
AX414431	1215 bp
	DNA
	linear
	PAT 14-JUN-2007

DEFINITION Sequence 1422 from Patent WO0228891.
ACCESSION AX414431
VERSION GI:21446888
KEYWORDS
SOURCE Listeria monocytogenes ATCC 19115.
ORANISM Listeria monocytogenes ATCC 19115.
REFERENCE 1 Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
AUTHORS Glaser, P. and Kunst, F.
TITLE Listeria innocua, genome and applications
JOURNAL Patent: WO 0228891-A 1422 11-Apr-2002;
Institut (FR)
FEATURES
source 1. .1215
/organism="Listeria monocytogenes ATCC 19115"
/db_xref="taxon:176281"
BASE COUNT 438 a 196 c 228 g 353 t
ORIGIN
Query Match 64.7%; Score 19.4; DB 6; Length 1215;
Best Local Similarity 95.2%; Pred. No. 6.4e+02;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 10 GGTAGACGGCGCTGATACCA 30
||||| |||||||
Db 923 GGTAGACGGCGCTGATACCA 943
RESULT 12
LMLISAPRF 3465 bp DNA linear BCT 31-JAN-1992
LOCUS LMLISAPRF
DEFINITION L. monocytogenes lisa gene for listeriolysin and part of prfa and
mpl genes.
ACCESSION X60035
VERSION X60035.1 GI:44110
KEYWORDS lisa gene; listeriolysin; mpl gene; prfa gene.
SOURCE Listeria monocytogenes.
ORANISM Listeria monocytogenes
REFERENCE 1 Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
AUTHORS Rasmussen, O. F.
TITLE Direct Submision
JOURNAL Submitted (04-JUN-1991) O. F. Rasmussen, Genetic Engineering Group,
Lundtoftevej 100, Bldg. 227, 2800 Lyngby, Denmark
REFERENCE 2 (bases 1 to 3465)
Rasmussen, O. F., Beck, T., Olsen, J. E., Dons, L. and Rossen, L.
TITLE Listeria monocytogenes isolates can be classified into two major
types according to the sequence of the listeriolysin gene
JOURNAL Infect. Immun. 59 (11), 3945-3951 (1991)
MEDLINE 92040062
PUBMED 1937753
FEATURES
source 1. .3465
/organism="Listeria monocytogenes"
/strain="12067"
/db_xref="taxon:1639"
complement(1. .277)
/gene="prfa"
complement(<1. .277)
/gene="prfa"
/codon_start=1
/transl_table=11
/protein_id="CAA42638.1"
/db_xref="GI:44111"
/db_xref="SPTREMBL:Q48789"
/translation="MSYNGDMTWLTGKPLAQTQMSLYQQLFAGIRYIDIRAKDNLKI
YHGPIFLNASSGVLETTQFLKKNPKETIIMRLKDEQNSNDSFDYRI"
misc_feature 278. .726
/note="intergen 1"
727. .2316
/gene="lisa"
727. .2316
/gene="lisa"
CDS
misc_feature
gene
CDS

/codon_start=1
/transl_table=11
/product="listeriolysin"
/protein_id="CAA42639.1"
/db_xref="GI:44112"
/db_xref="SWISS-PROT:P13128"
/translation="MKKIMLVETLILVSLPLAQTQMSLYQQLFAGIRYIDIRAKDNLKI
YHGPIFLNASSGVLETTQFLKKNPKETIIMRLKDEQNSNDSFDYRI"
PASPAPPIEKKHAEIDKTYDGNKNNVLYVGDVATVTPPKGKKDGEYIYVE
KKKSSINNNADIOVYNAISSLTIPGALVYKNSSEVENQPPVLPKRRSLTSLDPLG
MTNDNKTIVNATKSNVNNVNTLVERMNRVADYVNSAKIDYIDEMVISEQLI
AKFGTAFAVNNLSLVNNGAISSEKMOEVLSEFOIYVNVNNEPTFRSFGKAVTK
EQLQALGVNAENPPAYISSVAYGRQVYLLKLTNSHSTVYKAAFDAAVSGKSGVVEL
TNIKNSSEKAVIYGSAKDEVOIIDGNLGDRLDLKGAFFNRETPGVPLATYTNFL
KDNELAVIKNSSEYIETSKAYTDGKINHDSGGVYVAFQNFISMEINVDPGNEIYVH
KNMSENKSKLAHTSSYILDPGNARNINVAKECTGLAMEWRVYIDRNLPYKNNRN
ISIMETLIPKYSVDNPIE
2317. .2650
/note="intergen 2"
2651. .3465
/gene="mpl"
2651. .3465
/gene="mpl"
/codon_start=1
/transl_table=11
/protein_id="CAA42640.1"
/db_xref="GI:44113"
/db_xref="SWISS-PROT:P34025"
/translation="WKSRLICIMVIAFOAHFNNAVYKADSVGEERLRNNIQAKRNPA
LALPDSCEAKDFYKFKILMTKDKLGVHTYTLAISDGYLTNDDELKIVHTDNLK
AFINDLOOGLRITNOIKITEKNAIEKAPFALIGSEAHVKSIGNPYKEKEIINSR
TKRLVYVNIKLFAPEPEVASMTIOVDAEFGATILIKONMSEVERADTHKPDALGKAN
RLRPLRVHMINDLFYVDTHKSLITPDLNKTDSFQKVVSNKNTMTDP"
BASE COUNT 1246 a 569 c 647 g 1003 t
ORIGIN
Query Match 64.7%; Score 19.4; DB 1; Length 3465;
Best Local Similarity 95.2%; Pred. No. 5.3e+02;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 10 GGTAGACGGCGCTGATACCA 30
||||| |||||||
Db 3250 GGTAGACGGCGCTGATACCA 3270
RESULT 13
AA416781/c 4562 bp DNA linear PAT 14-JUN-2002
LOCUS AA416781
DEFINITION Sequence 3772 from Patent WO0228891.
ACCESSION AX416781
VERSION AX416781.1 GI:21445238
KEYWORDS
SOURCE Listeria monocytogenes ATCC 19115.
ORANISM Listeria monocytogenes ATCC 19115.
REFERENCE 1 Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
AUTHORS Glaser, P. and Kunst, F.
TITLE Listeria innocua, genome and applications
JOURNAL Patent: WO 0228891-A 3772 11-Apr-2002;
Pasteur Institut (FR)
FEATURES
source 1. .4562
/organism="Listeria monocytogenes ATCC 19115"
/db_xref="taxon:176281"
BASE COUNT 1166 a 943 c 782 g 1671 t
ORIGIN
Query Match 64.7%; Score 19.4; DB 6; Length 4562;
Best Local Similarity 95.2%; Pred. No. 5e+02;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 10 GGTAGACGGCGCTGATACCA 30
||||| |||||||
Db 3358 GGTAGACGGCGCTGATACCA 3338

LOCUS	DMMASTER	6333 bp	mRNA	linear	INV 02-MAY-1992
DEFINITION	D. melanogaster neurogenic locus mastermind mRNA for a nuclear protein.				
ACCESSION	X54251				
VERSION	X54251.1				
KEYWORDS	GT:8203				
SOURCE	mam gene; neurogenic.				
ORGANISM	fruit fly.				
	Drosophila melanogaster				
	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.				
REFERENCE	1 (bases 1 to 6333)				
AUTHORS	Yedvobnick,B.				
TITLE	Direct Submission				
JOURNAL	Submitted (08-AUG-1990) Yedvobnick B., Rollins Building, Emory University, Dept of Biology, 1510 Clifton Road, Atlanta Georgia 30322. U S A				
REFERENCE	2 (bases 1 to 6333)				
AUTHORS	Smoller,D., Friedel,C., Schmid,A., Bettler,D., Lam,L. and Yedvobnick,B.				
TITLE	The Drosophila neurogenic locus mastermind encodes a nuclear protein unusually rich in amino acid homopolymers				
JOURNAL	Genes Dev. 4 (10), 1688-1700 (1990)				
MEDLINE	91065516				
PUBMED	1701150				
FEATURES	Location/Qualifiers				

[illegible][illegible][illegible]

JOURNAL

Submitted (30-MAR-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Oct 14, 2001 this sequence version replaced gi:11094647.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: MAEV

Center clone name: RP23-313115

----- Summary Statistics

Sequencing vector: M13: L08821

Chemistry: Dye-primer Bodipy: 100% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 57463 bases at least Q40

Consensus quality: 76967 bases at least Q30

Consensus quality: 85864 bases at least Q20

Estimated insert size: 66691; sum-of-contigs estimation

Quality coverage: 0x in Q20 bases; agarose-gel estimation

----- Quality coverage: 0.7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see <http://www.hgsc.bcm.tmc.edu/docs/genbank.draft.data.html>).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 13 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

```
1      3582: contig of 3582 bp in length
*      3583      3682: gap of unknown length
*      3683      8017: contig of 4335 bp in length
*      8018      8117: gap of unknown length
*      8118      10383: contig of 2266 bp in length
*      10384      10483: gap of unknown length
*      10484      15321: contig of 4838 bp in length
*      15322      15421: gap of unknown length
*      15422      17484: contig of 2063 bp in length
*      17485      17585: gap of unknown length
*      17586      19771: contig of 2187 bp in length
*      19772      19871: gap of unknown length
*      19872      22127: contig of 2256 bp in length
*      22128      22227: gap of unknown length
*      22228      24233: contig of 2006 bp in length
*      24234      24333: gap of unknown length
*      24334      26451: contig of 2118 bp in length
*      26452      26551: gap of unknown length
*      26552      28764: contig of 2213 bp in length
*      28765      28864: gap of unknown length
*      28865      31103: contig of 2239 bp in length
*      31104      31203: gap of unknown length
*      31204      33261: contig of 2058 bp in length
*      33262      33361: gap of unknown length
*      33362      35372: contig of 2011 bp in length.
```

----- Location/Qualifiers

```
1..35372
  /organism="Mus musculus"
  /db_xref="taxon:10090"
  /chromosome="13"
  /clone="RP23-313115"
```

BASE COUNT 10592 a 6500 c 6766 g 10310 t 1204 others
ORIGIN

Query Match 64.7%: Score 19.4; DB 2: Length 35372;
Best Local Similarity 79.3%: Pred. No. 3.5e+02;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 2 ATGCCATGGGTAGACGAGGCTGATACCA 30
| ||| ||||| | ||||| | |

Db 9682 AGCCCTGGGTAGACAGGCTGATCCACA 9654

Search completed: November 23, 2002, 18:12:01
Job time : 2745 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 17:14:06 : Search time 257 Seconds
(without alignments)
262.879 Million cell updates/sec

Title: US-09-848-781-4

Perfect score: 30

Sequence: 1 catgcctagtgtagaacggctataccca 30

Scoring table: IDENTITY-NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

N_Geneseq_101002:*

1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	30	100.0	30	24	ABK12500
2	21	70.0	1533	21	ABO67931
3	21	70.0	1533	24	ABO69831
4	21	70.0	2944528	24	ABO63041
5	19.4	64.7	1215	24	ABO68609
6	19.4	64.7	4562	24	ABO70959
7	18.8	62.7	542	21	AAC74690
8	18.8	62.7	542	24	ABN17172
9	18.4	61.3	570	22	AAK38418

10	18	60.0	394	22	AAL16230	Human breast cancer
11	18	60.0	456	22	AAL20498	Human breast cancer
12	18	60.0	457	22	AAL11600	Human breast cancer
13	18	60.0	457	22	AAL23791	Human breast cancer
14	18	60.0	483	22	AAL14933	Human breast cancer
15	18	60.0	493	22	AAL25073	Human breast cancer
16	18	60.0	567	23	ABV57580	Human prostate cancer
17	18	60.0	143306	24	ABK49586	Human transporter
18	18	60.0	465237	24	ABO87681	Human oestrogen receptor
19	18	60.0	465237	24	ABO87681	Human oestrogen receptor
20	18	59.3	321	24	ABN94190	Gene #688 used to
21	17.8	59.3	581	21	AAAC59820	Human secreted protein
22	17.8	59.3	615	21	AAAC60834	Human secreted protein
23	17.8	59.3	641	21	AAAC60892	Human secreted protein
24	17.8	59.3	936	23	AAAC54139	Human secreted protein
25	17.8	59.3	1574	23	AAAC60836	Human secreted protein
26	17.8	59.3	2152	24	ABK63791	Human secreted protein
27	17.8	59.3	3711	24	ABK63791	Human secreted protein
28	17.8	59.3	4510	23	ABL20209	Drosophila melanogaster
29	17.8	59.3	4785	23	ABL13743	Drosophila melanogaster
30	17.8	59.3	18189	23	ABL16916	Drosophila melanogaster
31	17.8	59.3	22788	23	ABL20208	Drosophila melanogaster
32	17.8	59.3	49431	23	ABL13742	Drosophila melanogaster
33	17.8	59.3	249487	24	ABN85733	Mouse genomic region
34	17.4	58.0	424	23	ABV37697	Human prostate cancer
35	17.4	58.0	1592	21	AAAC57144	Human prostate cancer
36	17.4	58.0	4528	23	ABL30099	Drosophila melanogaster
37	17.4	58.0	7424	22	AAF28521	Drosophila melanogaster
38	17.4	58.0	7769	23	ABL30098	Drosophila melanogaster
39	17.4	58.0	11243	21	AAAB17338	N. meningitidis pathogen
40	17.4	58.0	19628	21	AAAB17338	N. meningitidis pathogen
41	17.4	58.0	119211	22	AAF28553	Human immune system
42	17.4	58.0	119211	22	AAF28553	Human immune system
43	17.2	57.3	608	22	AAAC3057	Neisseria meningitidis
44	17.2	57.3	615	24	AAAC3057	Human rhinovirus 1
45	17.2	57.3	615	24	ABK33106	DNA encoding novel

ALIGNMENTS

RESULT 1	
ID	ABK12500 standard; DNA: 30 BP.
XX	ABK12500;
AC	18-JUN-2002 (first entry)
XX	
DT	Listeria monocytogenes metalloprotease, PCR primer #1.
XX	
DE	Bacterial detection method: pathogen: food product; food retail;
XX	KW metalloprotease; gastroenteritis; food-related work area;
KW	mpi; PCR primer: ss.
XX	
OS	Listeria monocytogenes.
XX	
PN	WO200210433-A2.
XX	
PD	07-FEB-2002.
XX	
PF	03-MAY-2001; 2001WO-US14613.
XX	
PR	03-MAY-2000; 2000US-201405P.
XX	
PA	(EXPR-) EXPRESSIVE CONSTRUCTS INC.
XX	
PI	Sanders MC;
XX	
DR	WPI: 2002-303940/34.
XX	
PT	Detection of bacterial contamination in foods or food-related work
PT	areas comprises identifying a protein specific to the microorganism

PT using fluorescence or colorimetric methods -
XX
PS Disclosure; Page 13; 25pp; English.
XX
CC The invention describes a specific and sensitive method for detection of
CC a pathogenic microorganism in potentially contaminated food products
CC at the retail level, by detecting a bacteria-specific protein e.g. a
CC Listeria monocytogenes metalloprotease. The process is useful for
CC detecting Listeria monocytogenes, a common cause of gastroenteritis, or
CC other food contaminants in food products or food-related work areas.
CC This sequence represents the peptide M1 that functions as a substrate
CC for the Listeria monocytogenes metalloprotease mpl. The sequence
CC represents a primer used to clone Listeria monocytogenes metalloprotease
CC in order to determine the active or binding site of the protease and
CC increase the protease-substrate interaction used in the detection method.
XX
SQ Sequence 30 BP; 8 A; 8 C; 9 G; 5 T; 0 other;
Query Match 100.0%; Score 30; DB 24; Length 30;
Best Local Similarity 100.0%; Pred. No. 0.00015;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CATGCCATGGGTAGAACGGGCTGATACCA 30
DB 1 CATGCCATGGGTAGAACGGGCTGATACCA 30
RESULT 2
AB067931
ID AB067931 standard; DNM; 1533 BP.
XX
AC AB067931;
XX
DT 29-APR-2002 (first entry)
XX
DE Listeria monocytogenes EGD DNA sequence #55.
XX
KW Antibacterial; Listeria; food contamination; mutational analysis;
XX infection; ds.
OS Listeria monocytogenes EGD.
XX
XX WO200228891-A2.
XX
PD 11-APR-2002.
XX
PF 04-OCT-2001; 2001WO-FR03061.
XX
PR 04-OCT-2000; 2000FR-0012697.
XX
PA (INSP) INST PASTEUR.
XX (CNRS) CNRS CENT NAT RECH SCI.
XX
PI Kunst F, Glaeser P;
XX
DR WPI; 2002-332479/37.
XX
XX New genomic sequences from Listeria species, useful for detection,
PT treatment and prevention of infection, also related polypeptides,
PT antibodies and modulators -
XX
PS Claim 16; SEQ ID 744; 180pp; French.
XX
CC The present invention relates to nucleic acid sequences
CC (AB067188-AB071212) from Listeria sp. The sequences are useful as probes
CC and primers for identification and/or detection of Listeria (e.g. as
CC contaminants in foods, or mutational analysis) and for analysis of
CC gene expression. Proteins encoded by the nucleic acid sequences can be
CC used to screen for compounds that modulate gene expression, replication
CC and pathogenicity of Listeria (potential therapeutic agents), also for
CC treating infections by Listeria, and are useful as immunogens in
CC anti-Listeria vaccines.
CC Note: The sequence data for this patent did not form part

CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1533 BP; 532 A; 258 C; 331 G; 412 T; 0 other;
Query Match 70.0%; Score 21; DB 24; Length 1533;
Best Local Similarity 100.0%; Pred. No. 4.7;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 10 GGTAGAACGGGCTGATACCA 30
DB 600 GGTAGAACGGGCTGATACCA 620
RESULT 3
AB069831
ID AB069831 standard; DNM; 1533 BP.
XX
AC AB069831;
XX
DT 29-APR-2002 (first entry)
XX
DE Listeria monocytogenes EGDe DNA sequence #43.
XX
KW Antibacterial; Listeria; food contamination; mutational analysis;
XX infection; ds.
OS Listeria monocytogenes EGDe.
XX
XX WO200228891-A2.
XX
PD 11-APR-2002.
XX
PF 04-OCT-2001; 2001WO-FR03061.
XX
PR 04-OCT-2000; 2000FR-0012697.
XX
PA (INSP) INST PASTEUR.
XX (CNRS) CNRS CENT NAT RECH SCI.
XX
PI Kunst F, Glaeser P;
XX
DR WPI; 2002-332479/37.
XX
XX New genomic sequences from Listeria species, useful for detection,
PT treatment and prevention of infection, also related polypeptides,
PT antibodies and modulators -
XX
PS Claim 16; SEQ ID 2644; 180pp; French.
XX
CC The present invention relates to nucleic acid sequences
CC (AB067188-AB071212) from Listeria sp. The sequences are useful as probes
CC and primers for identification and/or detection of Listeria (e.g. as
CC contaminants in foods, or mutational analysis) and for analysis of
CC gene expression. Proteins encoded by the nucleic acid sequences can be
CC used to screen for compounds that modulate gene expression, replication
CC and pathogenicity of Listeria (potential therapeutic agents), also for
CC treating infections by Listeria, and are useful as immunogens in
CC anti-Listeria vaccines.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1533 BP; 532 A; 258 C; 331 G; 412 T; 0 other;
Query Match 70.0%; Score 21; DB 24; Length 1533;
Best Local Similarity 100.0%; Pred. No. 4.7;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 10 GGTAGAACGGGCTGATACCA 30
DB 600 GGTAGAACGGGCTGATACCA 620

```

RESULT 4
AB03041
ID   AB03041 standard; DNA; 2944528 BP.
XX
XX
AC   AB03041;
XX
XX
DT   05-FEB-2002 (first entry)
XX
XX
DE   Listeria monocytogenes EGD-e genome sequence.
XX
XX
KW   Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
XX   vitamin B12; bacterial infection; disease; ds.
XX
OS   Listeria monocytogenes.
XX
XX
PN   WO200177335-A2.
XX
XX
PD   18-OCT-2001.
XX
XX
PF   11-APR-2001; 2001WO-FR01118.
XX
XX
PR   11-APR-2000; 2000FR-0004629.
XX
XX
PA   (INSP ) INST PASTEUR.
XX
XX
PI   Buchrieser C, Frangeul L, Couve E, Rusniok C, Fsihi H, Dehoux P,
XX   Dussurget O, Chelouani F, Nedjati H, Glaser P, Kunst F, Cossart P,
XX   Daniels J, Goebel W, Kreft J, Kuhn M, Ng E, Vazquez-Boland JA;
XX   Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A;
XX   Charaborer T, Dommann E, Hain T, Berche P, Charbit A, Durant L,
XX   Perez-Piaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N;
XX   Madueno E, De Pablos B, Wehland J, Kaerst U, Entian K, Hauf J;
XX   Rose M, Voss H;
XX
XX
DR   WPI; 2002-010914/01.
XX
XX
PT   Genomic sequence for Listeria monocytogenes, useful e.g. for treatment
XX   and prevention of Listeria and related bacterial infections, and
XX   related polypeptides
XX
XX
PR   Claim 1; SEQ ID No 1; 192pp; French.
XX
XX
CC   The present sequence is the genome sequence of Listeria monocytogenes
XX   EGD-e. This sequence and fragments of this sequence are useful for
XX   selecting probes and primers for detecting genes in L. monocytogenes and
XX   related organisms, and to study genetic polymorphisms and other genomes.
XX   CC proteins (AB047297-AB050149) expressed from the present sequence are
XX   useful for raising specific antibodies, identification of L.
XX   CC monocytogenes and related organisms, and for biosynthesis and
XX   biodegradation, especially biosynthesis of Vitamin B12. This sequence and
XX   proteins encoded by it are also useful for selecting compounds that
XX   regulate gene expression and cell replication and modulate L.
XX   CC monocytogenes-related diseases. In addition, this sequence and proteins
XX   encoded by it are useful in pharmaceutical and vaccines compositions for
XX   the treatment or prevention of infections by L. monocytogenes and related
XX   organisms.
XX   CC Note: The sequence data for this patent did not form part of the printed
XX   CC specification, but was obtained in electronic format directly from WIPO
XX   CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX
SQ   Sequence 2944528 BP; 914202 A; 563301 C; 555061 G; 911964 T; 0 other;

Query Match      70.0%; Score 21; DB 24; Length 2944528;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY   10 GGTAGACGGCGCTGATACCA 30
DB 208338 GGTAGACGGCGCTGATACCA 208358

```

```

AB068609
ID   AB068609 standard; DNA; 1215 BP.
XX
XX
AC   AB068609;
XX
XX
DT   29-AUG-2002 (first entry)
XX
XX
DE   Listeria monocytogenes 4b contig DNA sequence #1375.
XX
XX
KW   Antibacterial; Listeria; food contamination; mutational analysis;
XX   infection; ds.
XX
OS   Listeria monocytogenes 4b.
XX
XX
PN   WO200228891-A2.
XX
XX
PD   11-APR-2002.
XX
XX
PF   04-OCT-2001; 2001WO-FR03061.
XX
XX
PR   04-OCT-2000; 2000FR-0012697.
XX
XX
PA   (INSP ) INST PASTEUR.
XX   (CNRS ) CNRS CENT NAT RECH SCI.
XX
XX
PI   Kunst F, Glaser P;
XX
XX
DR   WPI; 2002-332479/37.
XX
XX
PT   New genomic sequences from Listeria species, useful for detection,
XX   treatment and prevention of infection, also related polypeptides,
XX   antibodies and modulators
XX
XX
PR   Claim 14; SEQ ID 1422; 180pp; French.
XX
XX
CC   The present invention relates to nucleic acid sequences
XX   (AB067188-AB071212) from Listeria sp. The sequences are useful as probes
XX   and primers for identification and/or detection of Listeria (e.g. as
XX   contaminants in foods, or mutational analysis) and for analysis of
XX   CC gene expression. Proteins encoded by the nucleic acid sequences can be
XX   used to screen for compounds that modulate gene expression, replication
XX   CC and pathogenicity of Listeria (potential therapeutic agents), also for
XX   CC treating infections by Listeria, and are useful as immunogens in
XX   CC anti-Listeria vaccines.
XX   CC Note: The sequence data for this patent did not form part
XX   CC of the printed specification, but was obtained in electronic format
XX   CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
XX
SQ   Sequence 1215 BP; 438 A; 196 C; 228 G; 353 T; 0 other;

Query Match      64.7%; Score 19.4; DB 24; Length 1215;
Best Local Similarity 95.2%; Pred. No. 25;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY   10 GGTAGACGGCGCTGATACCA 30
DB 923 GGTAGACGGCGCTGATACCA 943

RESULT 6
AB070959/c
ID   AB070959 standard; DNA; 4562 BP.
XX
XX
AC   AB070959;
XX
XX
DT   29-AUG-2002 (first entry)
XX
XX
DE   Listeria monocytogenes 4b contig DNA sequence #901.
XX
XX
KW   Antibacterial; Listeria; food contamination; mutational analysis;
XX   infection; ds.
XX
OS   Listeria monocytogenes 4b.

```

XX WO200228891-A2.
XX 11-APR-2002.
XX
XX 04-OCT-2001; 2001WO-FR03061.
XX
XX 04-OCT-2000; 2000FR-0012697.
XX
XX (INSP) INST PASTEUR.
XX (CNRS) CNRS CENT NAT RECH SCI.
XX
XX Kunst F, Glaser P;
XX
XX WPI; 2002-332479/37.
XX
XX New genomic sequences from *Listeria* species, useful for detection,
XX treatment and prevention of infection, also related polypeptides,
XX antibodies and modulators -
XX
XX Claim 14; SEQ ID 3772; 180pp; French.
XX
XX The present invention relates to nucleic acid sequences
XX (AB067188-AB071212) from *Listeria* sp. The sequences are useful as probes
XX and primers for identification and/or detection of *Listeria* (e.g. as
XX contaminants in foods, or mutational analysis) and for analysis of
XX gene expression. Proteins encoded by the nucleic acid sequences can be
XX used to screen for compounds that modulate gene expression, replication
XX and pathogenicity of *Listeria* (potential therapeutic agents), also for
XX treating infections by *Listeria*, and are useful as immunogens in
XX anti-*Listeria* vaccines.
XX Note: The sequence data for this patent did not form part
XX of the printed specification, but was obtained in electronic format
XX directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 4562 BP; 1166 A; 943 C; 782 G; 1671 T; 0 other;
XX
Query Match 64.7%; Score 19.4; DB 24; Length 4562;
Best Local Similarity 95.2%; Pred. No. 31;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
QY 10 GGTAGAACGGCGTGATACCA 30
DB 3358 GGTAGACGGCGTGATACCA 3338
XX
RESULT 7
AAC74690/c
ID AAC74690 standard; cDNA; 542 BP.
XX
XX AAC74690;
XX
DT 08-FEB-2001 (first entry)
XX
XX Human ORFX ORF245 polynucleotide sequence SEQ ID NO:489.
XX
XX Human; open reading frame; ORFX; detection; cytosolic; hepatotropic;
XX vulnerable; antiparasitic; antiparkinsonian; neuroprotective;
XX anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
XX immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
XX hypotensive; dermatological; immunosuppressive; antiinflammatory;
XX antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
XX antianemic; gene therapy; cancer; proliferative disorder; hypertension;
XX neurodegenerative disorder; osteoarthritis; graft vs host disease;
XX cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
XX cholesterol ester storage; systemic lupus erythematosus; infection;
XX severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
XX allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
XX bone damage; cartilage damage; antiinflammatory disease; coagulation;
XX thrombosis; contraceptive; ss.
XX
XX Homo sapiens.
XX

PN WO200058473-A2.
XX 05-OCT-2000.
XX
XX 31-MAR-2000; 2000MO-US08621.
XX
XX 31-MAR-1999; 99US-0127607.
XX 02-APR-1999; 99US-0127636.
XX 05-APR-1999; 99US-0127728.
XX 30-MAR-2000; 2000US-0540763.
XX
XX (CURA-) CURAGEN CORP.
XX
XX Shimkets RA, Leach M;
XX
XX WPI; 2000-602362/57.
XX P-PsDB; AAB40481.
XX
XX Novel nucleic acids and peptides derived from open reading frame X,
XX useful for treating e.g. cancers, proliferative disorders,
XX neurodegenerative disorders and cardiovascular disease -
XX
XX Claim 5; Page 672; 5507pp; English.
XX
XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
XX which represent the human ORFX open reading frames 1 to 3161. The ORFX
XX sequences have activities such as: cytosolic; hepatotropic; vulnerable;
XX antiparasitic; antiparkinsonian; neuroprotective;
XX osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
XX immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
XX antidiabetic; hypotensive; dermatological; immunosuppressive;
XX antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
XX antithyroid; and antianemic. The sequences can be used for determining
XX the presence of or predisposition to, or preventing or treating
XX pathological conditions associated with an ORFX-associated disorder. The
XX nucleic acids can be used to express ORFX proteins in gene therapy
XX vectors. The proteins and nucleic acids may be used to treat cancers,
XX proliferative disorders, neurodegenerative disorders, osteoarthritis,
XX graft vs host disease, cardiovascular disease, diabetes mellitus,
XX hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
XX erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
XX bacterial or fungal infection, malaria, autoimmune disorders, asthma,
XX allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
XX nocturnal haemoglobinuria, antiinflammatory disease; to enhance
XX coagulation; to inhibit thrombosis; and as a contraceptive.
XX
SQ Sequence 542 BP; 101 A; 168 C; 170 G; 102 T; 1 other;
XX
Query Match 62.7%; Score 18.8; DB 21; Length 542;
Best Local Similarity 76.7%; Pred. No. 43;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
XX
QY 1 CATGCCATGGTAGAACGGCGTGATACCA 30
DB 177 CATGCCATGGCTGACCGCAGACACGA 148
XX
RESULT 8
ABN17172/c
ID ABN17172 standard; cDNA; 542 BP.
XX
XX ABN17172;
XX
DT 24-JUN-2002 (first entry)
XX
XX Human ORFX polynucleotide sequence SEQ ID NO:2821.
XX
XX Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;
XX hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
XX degenerative disorder; osteoarthritis; neurodegenerative disorder;
XX cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
XX hypertension; hypothyroidism; cholesterol ester storage disease;
XX immune deficiency; immune disorder; infectious disease;
XX

KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
KW myasthenia gravis; gene; ss.
XX
OS Homo sapiens.
XX WO200192523-A2.
XX
XX 06-DEC-2001.
XX
XX 29-MAY-2001; 2001WO-US10836.
XX
XX 30-MAY-2000; 2000US-206132P.
PR 29-AUG-2000; 2000US-228716P.
XX
XX (CURA-) CURAGEN CORP.
PA
XX Shinkets RA, Leach MD;
PI
XX WPI: 2002-106308/14.
DR P-PSDB: ABP01420.
XX
XX Novel human polypeptides and polynucleotides useful for diagnosing,
PT preventing and treating cardiovascular disease, neurodegenerative,
PT hyperproliferative disorders and autoimmune disorders -
XX
XX Disclosure; SEQ ID 2821: 1037pp; English.
XX
XX The present invention describes substantially purified human proteins
CC (referred to as open reading frame, ORFX, where x is 1-11491 (see Table 1
CC in the specification). ABN15762 to ABN27252 encode the human ORFX
CC proteins given in ABP00010 to ABP11500. ORFX proteins are useful for
CC treating or preventing a pathology associated with an ORFX-associated
CC disorder in humans, and in the manufacture of a medicament for treating a
CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide
CC sequences can be used in gene therapy. ORFX sequences can be used in the
CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
CC psoriasis, benign tumors, keloid, degenerative disorders, hemorrhage,
CC osteoarthritis, neurodegenerative disorders, disorders related to organ
CC transplantation, cardiovascular diseases, diabetes mellitus, systemic
CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester
CC storage disease, various immune deficiencies and disorders, infectious
CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
CC disease and autoimmune inflammatory eye disease. ORFX proteins are also
CC useful for treating burns, incisions, ulcers, for treating osteoporosis,
CC bone degenerative disorders, or periodontal disease, and for gut
CC protection or regeneration and treatment of lung or liver fibrosis,
CC reperfusion injury in various tissues and conditions resulting from
CC systemic cytokine damage.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIFO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 542 BP; 101 A; 168 C; 170 G; 102 T; 1 other;

Query Match 62.7%; Score 18.8; DB 24; Length 542;
Best Local Similarity 76.7%; Pred. No. 43;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CATGCCATGGGTAGAACGGGCTGATACCA 30
IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII
DB 177 CATGCCATGGGTAGAACGGGCTGATACCA 148

RESULT 9
AAK38418
ID AAK38418 standard; DNA; 570 BP.
XX
XX AAK38418;
XX
XX 06-NOV-2001 (first entry)
XX
XX Human bone marrow expressed single exon probe SEQ ID NO: 12975.

XX
KW Human: bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
OS Homo sapiens.
XX WO200157276-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US00668.
XX
XX 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI
XX WPI: 2001-488900/53.
DR
XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human bone marrow -
XX
XX Example 4; SEQ ID NO: 12975; 658pp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukaemia and myeloma. The present sequence is one of
CC the probes of the invention.
XX
SQ Sequence 570 BP; 194 A; 132 C; 97 G; 147 T; 0 other;

Query Match 61.3%; Score 18.4; DB 22; Length 570;
Best Local Similarity 78.6%; Pred. No. 66;
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 ATGCCATGGGTAGAACGGGCTGATACCC 29
IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII
DB 433 ATGCCATGGGTAGAACGGGCTGATACCC 460

RESULT 10
AAL16230
ID AAL16230 standard; CDNA; 394 BP.
XX
XX AAL16230;
XX
XX 07-DEC-2001 (first entry)
XX
XX Human breast cancer expressed polynucleotide 8687.
XX
XX Human: breast cancer; cell marker; cytosolic; ss.
XX
OS Homo sapiens.
XX WO200151628-A2.
XX
XX 19-JUL-2001.
XX
XX 10-JAN-2001; 2001WO-US00798.
XX
XX 14-JAN-2000; 2000US-0176077.
PR 14-MAR-2000; 2000US-0189167.
PR 24-MAR-2000; 2000US-0192099.
PR 29-MAR-2000; 2000US-0193480.

PR 15-MAY-2000; 2000US-0205230.
PR 09-JUN-2000; 2000US-0211315.
PR 25-JUL-2000; 2000US-0220534.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
PI Lillie J, Xu Y, Wang Y, Steinmann K;
XX
DR WPI; 2001-451856/48.
XX
PT New peptide useful as a marker for the diagnosis of breast cancer -
XX
PS Claim 1; Page 1572; 3695pp; English.
XX
CC The invention relates to human breast cancer expressed polynucleotides
CC (AAL07544-AAL26789) and methods of assessing whether a patient is
CC afflicted with breast cancer by examining the correlation between the
CC expression of certain markers and the cancerous state of breast cells.
CC The polynucleotides and encoded polypeptides are potential markers for
CC detecting, diagnosing, monitoring, characterizing treating and
CC potentially preventing breast cancer. The polynucleotides and encoded
CC polypeptides are also useful for isolating compounds with cytostatic
CC activity.
XX
SQ Sequence 394 BP; 96 A; 90 C; 115 G; 93 T; 0 other;
XX
Query Match 60.0%; Score 18; DB 22; Length 394;
Best Local Similarity 80.8%; Pred. No. 96;
Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
XX
OY 3 TGCCATGGGTAGAACGGCGTGATACC 28
DB 206 TGCCAAAGGTAAATGGGCGTTACC 231
XX
RESULT 11
AAL20498
ID AAL20498 standard; cDNA; 456 BP.
XX
AC AAL20498;
XX
DT 07-DEC-2001 (first entry)
XX
DE Human breast cancer expressed polynucleotide 12955.
XX
KW Human; breast cancer; cell marker; cytostatic; ss.
XX
OS Homo sapiens.
XX
PN WO200151628-A2.
XX
PD 19-JUL-2001.
XX
PE 10-JAN-2001; 2001WO-US00798.
XX
PR 14-JAN-2000; 2000US-0176077.
PR 14-MAR-2000; 2000US-0189167.
PR 24-MAR-2000; 2000US-0192099.
PR 29-MAR-2000; 2000US-0193480.
PR 15-MAY-2000; 2000US-0205230.
PR 09-JUN-2000; 2000US-0211315.
PR 25-JUL-2000; 2000US-0220534.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
PI Lillie J, Xu Y, Wang Y, Steinmann K;
XX
DR WPI; 2001-451856/48.
XX
PT New peptide useful as a marker for the diagnosis of breast cancer -
XX
PS Claim 1; Page 2294; 3695pp; English.
XX

CC The invention relates to human breast cancer expressed polynucleotides
CC (AAL07544-AAL26789) and methods of assessing whether a patient is
CC afflicted with breast cancer by examining the correlation between the
CC expression of certain markers and the cancerous state of breast cells.
CC The polynucleotides and encoded polypeptides are potential markers for
CC detecting, diagnosing, monitoring, characterizing treating and
CC potentially preventing breast cancer. The polynucleotides and encoded
CC polypeptides are also useful for isolating compounds with cytostatic
CC activity.
XX
SQ Sequence 456 BP; 116 A; 99 C; 132 G; 109 T; 0 other;
XX
Query Match 60.0%; Score 18; DB 22; Length 456;
Best Local Similarity 80.8%; Pred. No. 98;
Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
XX
OY 3 TGCCATGGGTAGAACGGCGTGATACC 28
DB 183 TGCCAAAGGTAAATGGGCGTTACC 208
XX
RESULT 12
AAL11600
ID AAL11600 standard; cDNA; 457 BP.
XX
AC AAL11600;
XX
DT 07-DEC-2001 (first entry)
XX
DE Human breast cancer expressed polynucleotide 4057.
XX
KW Human; breast cancer; cell marker; cytostatic; ss.
XX
OS Homo sapiens.
XX
PN WO200151628-A2.
XX
PD 19-JUL-2001.
XX
PE 10-JAN-2001; 2001WO-US00798.
XX
PR 14-JAN-2000; 2000US-0176077.
PR 14-MAR-2000; 2000US-0189167.
PR 24-MAR-2000; 2000US-0192099.
PR 29-MAR-2000; 2000US-0193480.
PR 15-MAY-2000; 2000US-0205230.
PR 09-JUN-2000; 2000US-0211315.
PR 25-JUL-2000; 2000US-0220534.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
PI Lillie J, Xu Y, Wang Y, Steinmann K;
XX
DR WPI; 2001-451856/48.
XX
PT New peptide useful as a marker for the diagnosis of breast cancer -
XX
PS Claim 1; Page 740; 3695pp; English.
XX
CC The invention relates to human breast cancer expressed polynucleotides
CC (AAL07544-AAL26789) and methods of assessing whether a patient is
CC afflicted with breast cancer by examining the correlation between the
CC expression of certain markers and the cancerous state of breast cells.
CC The polynucleotides and encoded polypeptides are potential markers for
CC detecting, diagnosing, monitoring, characterizing treating and
CC potentially preventing breast cancer. The polynucleotides and encoded
CC polypeptides are also useful for isolating compounds with cytostatic
CC activity.
XX
SQ Sequence 457 BP; 116 A; 100 C; 131 G; 109 T; 1 other;
XX
Query Match 60.0%; Score 18; DB 22; Length 457;
Best Local Similarity 80.8%; Pred. No. 98;
XX

Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 3 TGCCATGGGTAGAACGGCTGATACC 28
||||| ||||| ||||| |||||
Db 185 TGCCAAGGGTAAATGGGCTGTACC 210

RESULT 13
AAL23791/C
ID AAL23791 standard; cDNA: 457 BP.

XX AAL23791;

DT 07-DEC-2001 (first entry)

DE Human breast cancer expressed polynucleotide 16248.

KW Human; breast cancer; cell marker; cytostatic; ss.

OS Homo sapiens.

XX WO200151628-A2.

PD 19-JUL-2001.

PE 10-JAN-2001; 2001WO-US00798.

PR 14-JAN-2000; 2000US-0176077.

PR 14-MAR-2000; 2000US-0189167.

PR 24-MAR-2000; 2000US-0192099.

PR 29-MAR-2000; 2000US-0193480.

PR 15-MAY-2000; 2000US-0205230.

PR 09-JUN-2000; 2000US-0211315.

PR 25-JUL-2000; 2000US-0220534.

PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

PI Lillie J, Xu Y, Wang Y, Steinmann K;

DR WPI: 2001-451856/48.

PT New peptide useful as a marker for the diagnosis of breast cancer -

XX Claim 1; Page 2969-2970; 3695pp; English.

XX The invention relates to human breast cancer expressed polynucleotides

CC (AAL07544-AAL26789) and methods of assessing whether a patient is

CC afflicted with breast cancer by examining the correlation between the

CC expression of certain markers and the cancerous state of breast cells.

CC The polynucleotides and encoded polypeptides are potential markers for

CC detecting, diagnosing, monitoring, characterizing treating and

CC potentially preventing breast cancer. The polynucleotides and encoded

CC polypeptides are also useful for isolating compounds with cytostatic

CC activity.

XX Sequence 457 BP; 111 A; 133 C; 97 G; 116 T; 0 other;

SO Query Match 60.0%; Score 18; DB 22; Length 457;

XX Best Local Similarity 80.8%; Pred. No. 98;

XX Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 3 TGCCATGGGTAGAACGGCTGATACC 28
||||| ||||| ||||| |||||
Db 275 TGCCAAGGGTAAATGGGCTGTACC 250

RESULT 14

AAL14933/C

ID AAL14933 standard; cDNA: 483 BP.

XX AAL14933;

DT 07-DEC-2001 (first entry)

XX Human breast cancer expressed polynucleotide 7390.

DE Human; breast cancer; cell marker; cytostatic; ss.

KW Homo sapiens.

OS WO200151628-A2.

XX 19-JUL-2001.

PD 10-JAN-2001; 2001WO-US00798.

PE 14-JAN-2000; 2000US-0176077.

PR 14-MAR-2000; 2000US-0189167.

PR 24-MAR-2000; 2000US-0192099.

PR 29-MAR-2000; 2000US-0193480.

PR 15-MAY-2000; 2000US-0205230.

PR 09-JUN-2000; 2000US-0211315.

PR 25-JUL-2000; 2000US-0220534.

PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

PI Lillie J, Xu Y, Wang Y, Steinmann K;

DR WPI: 2001-451856/48.

PT New peptide useful as a marker for the diagnosis of breast cancer -

XX Claim 1; Page 1331; 3695pp; English.

XX The invention relates to human breast cancer expressed polynucleotides

CC (AAL07544-AAL26789) and methods of assessing whether a patient is

CC afflicted with breast cancer by examining the correlation between the

CC expression of certain markers and the cancerous state of breast cells.

CC The polynucleotides and encoded polypeptides are potential markers for

CC detecting, diagnosing, monitoring, characterizing treating and

CC potentially preventing breast cancer. The polynucleotides and encoded

CC polypeptides are also useful for isolating compounds with cytostatic

CC activity.

XX Sequence 483 BP; 114 A; 142 C; 107 G; 120 T; 0 other;

SO Query Match 60.0%; Score 18; DB 22; Length 483;

XX Best Local Similarity 80.8%; Pred. No. 99;

XX Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 3 TGCCATGGGTAGAACGGCTGATACC 28
||||| ||||| ||||| |||||
Db 301 TGCCAAGGGTAAATGGGCTGTACC 276

RESULT 15

AAL25073

ID AAL25073 standard; cDNA: 493 BP.

XX AAL25073;

DT 07-DEC-2001 (first entry)

DE Human breast cancer expressed polynucleotide 17530.

KW Human; breast cancer; cell marker; cytostatic; ss.

OS Homo sapiens.

XX WO200151628-A2.

XX 19-JUL-2001.

PD 10-JAN-2001; 2001WO-US00798.

PE 14-JAN-2000; 2000US-0176077.

PR 14-MAR-2000: 2000US-0189167.
 PR 24-MAR-2000: 2000US-0192059.
 PR 29-MAR-2000: 2000US-0193480.
 PR 15-MAY-2000: 2000US-0205230.
 PR 09-JUN-2000: 2000US-0211315.
 PR 25-JUL-2000: 2000US-0220534.
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX
 XX
 PI lllle J, Xu Y, Wang Y, Steimann K;
 XX
 DR WPI, 2001-451856/48.
 XX
 PT New peptide useful as a marker for the diagnosis of breast cancer -
 XX
 PS Claim 1; Page 3240; 3695pp; English.
 XX
 CC The invention relates to human breast cancer expressed polynucleotides
 CC (AAU07544-AAU26789) and methods of assessing whether a patient is
 CC afflicted with breast cancer by examining the correlation between the
 CC expression of certain markers and the cancerous state of breast cells.
 CC The polynucleotides and encoded polypeptides are potential markers for
 CC detecting, diagnosing, monitoring, characterising, treating and
 CC potentially preventing breast cancer. The polynucleotides and encoded
 CC polypeptides are also useful for isolating compounds with cytostatic
 CC activity.
 CC
 SQ Sequence 493 BP; 125 A; 110 C; 141 G; 117 T; 0 other;
 XX
 QY Query Match 60.0%; Score 18; DB 22; Length 493;
 Best Local Similarity 80.8%; Pred. No. 99;
 Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 Db 3 TGGCATGGGTAGAGGCGCTGATACC 28
 179 TGGCAAGGCTAAATGGGCGCTGTACC 204

Search completed: November 23, 2002, 17:34:27
 Job time : 770 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 17:20:41 : Search time 50 Seconds
(without alignments)
184.006 Million cell updates/sec

Title: US-09-848-781-4
Perfect score: 30
Sequence: 1 catgcatgggtagaacggctgtaccaca 30

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued_patents_NA: *
1: /cgn2_6/ptodata/1/lna/5A_COMB.seq: *
2: /cgn2_6/ptodata/1/lna/5B_COMB.seq: *
3: /cgn2_6/ptodata/1/lna/6A_COMB.seq: *
4: /cgn2_6/ptodata/1/lna/6B_COMB.seq: *
5: /cgn2_6/ptodata/1/lna/PTCUTS_COMB.seq: *
6: /cgn2_6/ptodata/1/lna/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	17.8	59.3	615	US-09-189-527-6	Sequence 6, Appli
2	17.8	59.3	641	US-09-615-192A-185	Sequence 185, App
3	17.8	59.3	1574	US-09-189-527-10	Sequence 10, Appli
4	17.8	59.3	35060	US-08-814-095-7	Sequence 7, Appli
5	17.2	57.3	627	US-08-221-816B-31	Sequence 31, Appli
6	16.8	56.0	465	US-08-300-928C-9	Sequence 9, Appli
7	16.8	56.0	465	US-08-430-944D-9	Sequence 9, Appli
8	16.8	56.0	465	US-08-431-184-9	Sequence 9, Appli
9	16.8	56.0	465	US-08-300-928C-7	Sequence 7, Appli
10	16.8	56.0	465	US-08-430-944D-7	Sequence 7, Appli
11	16.8	56.0	465	US-08-430-944D-7	Sequence 7, Appli
12	16.8	56.0	465	US-08-431-184-7	Sequence 7, Appli
13	16.8	56.0	476	US-08-300-928C-5	Sequence 5, Appli
14	16.8	56.0	476	US-08-430-944D-5	Sequence 5, Appli
15	16.8	56.0	476	US-08-430-014-5	Sequence 5, Appli
16	16.8	56.0	476	US-08-431-184-5	Sequence 5, Appli
17	16.8	56.0	485	US-07-807-529A-5	Sequence 5, Appli
18	16.8	56.0	485	PCT-US93-02462-5	Sequence 5, Appli
19	16.8	56.0	2870	US-09-221-017B-205	Sequence 205, App
20	16.8	56.0	85	US-09-364-543-10	Sequence 10, Appli
21	16.4	54.7	86	US-09-364-543-11	Sequence 11, Appli
22	16.4	54.7	3839	US-09-245-248B-57	Sequence 57, Appli
23	16.4	54.0	694	US-08-998-416-1102	Sequence 1102, Ap
24	16.2	54.0	933	US-08-612-857-1	Sequence 1, Appli
25	16.2	54.0	949	US-08-167-035-32	Sequence 32, Appli
26	16.2	54.0	949	US-08-167-035-32	Sequence 32, Appli
27	16.2	54.0	949	US-08-167-035-32	Sequence 50, Appli

28	16.2	54.0	949	US-08-208-887A-32	Sequence 32, Appli
29	16.2	54.0	949	US-08-539-005-32	Sequence 32, Appli
30	16.2	54.0	949	US-08-539-005-32	Sequence 50, Appli
31	16.2	54.0	949	US-09-280-598-34	Sequence 34, Appli
32	16.2	54.0	1072	US-07-906-349A-2	Sequence 2, Appli
33	16.2	54.0	1072	US-08-167-035-5	Sequence 5, Appli
34	16.2	54.0	1072	US-08-167-035-5	Sequence 49, Appli
35	16.2	54.0	1072	US-08-208-887A-5	Sequence 5, Appli
36	16.2	54.0	1072	US-08-539-005-5	Sequence 5, Appli
37	16.2	54.0	1072	US-08-539-005-49	Sequence 49, Appli
38	16.2	54.0	1072	US-09-280-598-2	Sequence 2, Appli
39	16.2	54.0	1109	US-08-664-962B-5	Sequence 5, Appli
40	16.2	54.0	1109	US-09-311-743-5	Sequence 5, Appli
41	16.2	54.0	1131	US-09-410-464-7	Sequence 7, Appli
42	16.2	54.0	1308	US-09-410-464-6	Sequence 6, Appli
43	16.2	54.0	1660	US-09-297-468-1	Sequence 1, Appli
44	16.2	54.0	5656	US-09-410-464-5	Sequence 5, Appli
45	16	53.3	732	US-09-336-536-75	Sequence 75, Appli

ALIGNMENTS

```
RESULT 1
US-09-189-527-6/c
Sequence 6, Application US/09189527A
Patent No. 6387639
GENERAL INFORMATION:
APPLICANT: Jerome B. Posner
APPLICANT: Joseph O. Dalmau
APPLICANT: Myrna R. Rosenfeld
TITLE OF INVENTION: Ma Family Polypeptides and Anti-Ma
FILE REFERENCE: SLK98-01
CURRENT APPLICATION NUMBER: US/09/189,527A
CURRENT FILING DATE: 1998-11-10
NUMBER OF SEQ ID NOS: 14
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 6
LENGTH: 615
TYPE: DNA
ORGANISM: homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (1)...(585)
US-09-189-527-6
Query Match 59.3%; Score 17.8; DB 4; Length 615;
Best Local Similarity 75.9%; Pred. No. 17;
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
OY 2 ATGCCATGGGTAGAACGGGCTGATACCA 30
DB 481 ATCTCATGGGTAGCAGGGCTGATAGCGCA 453
RESULT 2
US-09-615-192A-185
Sequence 185, Application US/09615192A
Patent No. 6410718
GENERAL INFORMATION:
APPLICANT: Bloksberg, Leonard N.
APPLICANT: Havukkala, Ilkka
TITLE OF INVENTION: Materials and Methods for the
FILE REFERENCE: 11000.1003c4U
CURRENT APPLICATION NUMBER: US/09/615,192A
CURRENT FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 08/975,316
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: US 08/713,000
PRIOR FILING DATE: 1996-09-11
PRIOR APPLICATION NUMBER: US 09/169,789
```

PRIOR FILING DATE: 1998-10-09
 NUMBER OF SEQ ID NOS: 405
 SOFTWARE: FastSeq for Windows Version 3.0
 SEQ ID NO 185
 LENGTH: 641
 TYPE: DNA
 ORGANISM: Pinus radiata
 US-09-615-192A-185

Query Match 59.3%; Score 17.8; DB 4; Length 641;
 Best Local Similarity 75.9%; Pred. No. 17;
 Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CATGCCATGGTAGAAGGGCTGATACCC 29
 DB 535 CCTGCCTTGGATGACGCTGTACAC 563

RESULT 3
 US-09-189-527-10/c
 Sequence 10, Application US/09189527A
 Patent No. 6387639

GENERAL INFORMATION:
 APPLICANT: Jerome B. Posner
 APPLICANT: Josep O. Dalmau
 APPLICANT: Myrna R. Rosenfeld
 TITLE OF INVENTION: Ma Family Polypeptides and Anti-Ma
 TITLE OF INVENTION: Antibodies
 FILE REFERENCE: SLK98-01
 CURRENT APPLICATION NUMBER: US/09/189,527A
 CURRENT FILING DATE: 1998-11-10
 NUMBER OF SEQ ID NOS: 14
 SOFTWARE: FastSeq for Windows Version 3.0
 SEQ ID NO 10
 LENGTH: 1574
 TYPE: DNA
 ORGANISM: homo sapiens
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (1)...(850)
 US-09-189-527-10

Query Match 59.3%; Score 17.8; DB 4; Length 1574;
 Best Local Similarity 75.9%; Pred. No. 20;
 Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 ATGCCATGGCTGACAGGGCTGATACCA 30
 DB 242 ATCTCATGGGTAGCAGGGCTGAGCGCA 214

RESULT 4
 US-08-814-095-7/c
 Sequence 7, Application US/08814095
 Patent No. 6025183
 GENERAL INFORMATION:
 APPLICANT: Soreq, Hermona
 APPLICANT: Zakut, Haim
 APPLICANT: Shaul, Moshe
 TITLE OF INVENTION: TRANSGENIC ANIMAL ASSAY SYSTEM FOR
 TITLE OF INVENTION: ANTI-CHOLINESTERASE SUBSTANCES
 NUMBER OF SEQUENCES: 7
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: KOHN & ASSOCIATES
 STREET: 30500 NO. 6025183thwestern Highway, Suite 410
 CITY: Farmington Hills
 STATE: Michigan
 COUNTRY: U.S.
 ZIP: 48334
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/814,095
 FILING DATE:
 CLASSIFICATION: 800
 ATTORNEY/AGENT INFORMATION:
 NAME: Montgomery, Ilene N.
 REGISTRATION NUMBER: 38,972
 REFERENCE/DOCKET NUMBER: 2391.00066
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (248) 539-5050
 TELEFAX: (248) 539-5055
 INFORMATION FOR SEQ ID NO: 7:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 35060 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: other nucleic acid
 DESCRIPTION: /desc = "Cosmid including ACHE
 DESCRIPTION: promoter, ACHE gene and ARS gene"
 HYPOTHEICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 ORGANISM: Homo sapiens
 POSITION IN GENOME:
 CHROMOSOME/SEGMENT: 7q22
 FEATURE:
 NAME/KEY: promoter
 LOCATION: 4089..22464
 OTHER INFORMATION: /function= "ACHE Promotor"
 OTHER INFORMATION: /standard_name= "ACHE Promotor"
 FEATURE:
 NAME/KEY: exon
 LOCATION: 22465..22537
 OTHER INFORMATION: /function= "non-translated"
 OTHER INFORMATION: /gene= "ACHE"
 OTHER INFORMATION: /number= 1
 FEATURE:
 NAME/KEY: exon
 LOCATION: 24090..25177
 IDENTIFICATION METHOD: experimental
 OTHER INFORMATION: /function= "(translation start:
 OTHER INFORMATION: 24110)"
 OTHER INFORMATION: /evidence= EXPERIMENTAL
 OTHER INFORMATION: /gene= "ACHE"
 OTHER INFORMATION: /number= 2
 FEATURE:
 NAME/KEY: exon
 LOCATION: 2524..26009
 IDENTIFICATION METHOD: experimental
 OTHER INFORMATION: /evidence= EXPERIMENTAL
 OTHER INFORMATION: /gene= "ACHE"
 OTHER INFORMATION: /number= 3
 FEATURE:
 NAME/KEY: exon
 LOCATION: 27005..27274
 IDENTIFICATION METHOD: experimental
 OTHER INFORMATION: /evidence= EXPERIMENTAL
 OTHER INFORMATION: /gene= "ACHE"
 OTHER INFORMATION: /number= 4
 FEATURE:
 NAME/KEY: exon
 LOCATION: 27255..28007
 IDENTIFICATION METHOD: experimental
 OTHER INFORMATION: /evidence= EXPERIMENTAL
 OTHER INFORMATION: /gene= "ACHE"
 OTHER INFORMATION: /number= 5
 FEATURE:
 NAME/KEY: terminator
 LOCATION: 27385..27387
 FEATURE:
 NAME/KEY: exon

LOCATION: 2808..28129
IDENTIFICATION METHOD: experimental
OTHER INFORMATION: /evidence= EXPERIMENTAL
OTHER INFORMATION: /gene= "ACHE"
OTHER INFORMATION: /number= 6
FEATURE:
NAME/KEY: terminator
LOCATION: 28129..28131
FEATURE:
NAME/KEY: exon
LOCATION: complement (34528..34895)
OTHER INFORMATION: /function= "arsenite resistance
OTHER INFORMATION: gene"
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 1
FEATURE:
NAME/KEY: exon
LOCATION: complement (34092..34358)
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 2
FEATURE:
NAME/KEY: exon
LOCATION: complement (33779..33963)
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 3
FEATURE:
NAME/KEY: exon
LOCATION: complement (33493..33591)
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 4
FEATURE:
NAME/KEY: exon
LOCATION: complement (33297..33408)
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 5
FEATURE:
NAME/KEY: exon
LOCATION: complement (32959..33094)
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 6
FEATURE:
NAME/KEY: exon
LOCATION: complement (32569..32628)
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 7
FEATURE:
NAME/KEY: exon
LOCATION: complement (32386..32468)
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 8
FEATURE:
NAME/KEY: exon
LOCATION: complement (31894..32080)
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 9
FEATURE:
NAME/KEY: exon
LOCATION: complement (31363..31534)
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 10
FEATURE:
NAME/KEY: exon
LOCATION: complement (31131..31284)
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 11
FEATURE:
NAME/KEY: exon
LOCATION: complement (30816..31011)
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 12
FEATURE:
NAME/KEY: exon
LOCATION: complement (30470..30626)

OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 13
FEATURE:
NAME/KEY: exon
LOCATION: complement (30187..30274)
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 14
FEATURE:
NAME/KEY: exon
LOCATION: complement (29945..30073)
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 15
FEATURE:
NAME/KEY: exon
LOCATION: complement (29664..29856)
OTHER INFORMATION: /gene= "ARS"
OTHER INFORMATION: /number= 16
US-08-814-095-7
Query Match 59.3%; Score 17.8; DB 3; Length 35060;
Best Local Similarity 90.5%; Pred. No. 37;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Oy 8 TGGGTAGAAGCGGCTGATACC 28
||||||| ||||| |||
Db 6086 TGGGTAGAATGGCTGAGACC 6066
RESULT 5
US-08-221-816B-31/C
Sequence 31, Application US/08221816B
Patent No. 5738985
GENERAL INFORMATION:
APPLICANT: Miles, Vincent J.
APPLICANT: Mathews, Michael B.
APPLICANT: Katze, Michael G.
APPLICANT: Witherell, Gary
APPLICANT: Watson, Julia C.
TITLE OF INVENTION: METHOD FOR SELECTIVE INACTIVATION
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036/2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/221.816B
FILING DATE: 01-APR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7960-030
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 627 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: RNA
US-08-221-816B-31

Query Match	57.3%;	Score 17.2;	DB 1;	Length 627;
Best Local Similarity	73.3%;	Pred. No. 32;		
Matches 22; Conservative	0;	Mismatches 8;	Indels 0;	Gaps 0;

```

QY      1  CATGCCATGGGTAGAACGGGCTGATACCA  30
          ||| | | | | | | | | | | | | |
Db      43  CACCCATGGGTGAATGGTGGGATACCA  14

```

RESULT 6
US-08-300-928C-9

```

; GENERAL INFORMATION:
; APPLICANT: GEFTER, Malcolm L. et al.

```

```

; NUMBER OF SEQUENCES: 101
;
; CORRESPONDENCE ADDRESS:

```

;
;
; CITY: WALTHAM
; STATE: Massachusetts

```

;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk

```

```

; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:

```

```

CLASSIFICATION: 435
PRIOR APPLICATION DATA:

```

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:

REFERENCE/DOCKET NUMBER: 002.6US(IMI-044)

```

: 1EEDRA. (011) 22/ 0541
: INFORMATION FOR SEQ ID NO: 9
: SEQUENCE CHARACTERISTICS:

```

LIFE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear

FEATURE:	
NAME/KEY:	CDS
LOCATION:	1 291

Query Match	56.0%;	Score 16.8;	DB 3;	Length 465;
Best Local Similarity	75.0%;	Pred No 47.		

QY 3 TGGCATGGGTAGAACGGGCTGATACCA 30
|||||

RECETTE 7

; Sequence 9, Application US/08430944D
; Patent No. 6025162

```

: TITLE OF INVENTION:  A HUMAN T CELL REACTIVE FELINE PROTEIN
:
: TITLE OF INVENTION:
:
: NUMBER OF SEQUENCES:  103
:
: CORRESPONDENCE ADDRESS:
:
:

```

STREET: 28 State Street
CITY: Boston
STATE: Massachusetts

```

;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy

```

```

; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:

```

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/430,014

APPLICATION NUMBER: US 08/300,928
FILING DATE: 02-SEPT-1994

REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: IMI-044DV2

TELEPHONE: (617) 221-7400
TELEFAX: (617) 742-4214
INFORMATION FOR SEO ID NO: 9:

STRANDEDNESS: single

NAME/KEY.	CDS
FEATURE:	
MODECODE 11F.	CJNNA

```

; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 58_291

```

Query Match	56.0%;	Score 16.8;	DB 3;	Length 465;
Best local similarity	75.0%;			
Best local alignment	75.0%;	Prod No. 47.		

3 TGCCATGGGTAGAACGGGCTGATACCA 30

DESCRITTORE

Sequence 9, Application US/08430014
Patent No. 6048962
GENERAL INFORMATION:

```

1      TITLE OF INVENTION: PEPTIDES FOR HUMAN T CELL REACTIVE FELINE
2      TITLE OF INVENTION: PROTEIN (TRFP)
3      NUMBER OF SEQUENCES: 101

```

ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION
STREET: 610 LINCOLN STREET

```

; COUNTRY: USA
; ZIP: 02145

```

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/430,014
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/300,928
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: AMY E. MANDRAGOURAS
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: 002.60S(IMI-044)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 465 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..291
US-08-430-014-9

Query Match 56.0%; Score 16.8; DB 3; Length 465;
Best Local Similarity 75.0%; Pred. No. 47;
Matches 21; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 3 TGGCATGGGTAGAACGGCTGATACCCA 30
||| | | | | | | | | | | | | | |
Db 199 TGCTACGTGAGAACGCGACTCATATCCA 226

RESULT 9
US-08-431-184-9
Sequence 9, Application US/08431184
Patent No. 6120769
GENERAL INFORMATION:
APPLICANT: Bruce L. Rogers et al.
TITLE OF INVENTION: A HUMAN T CELL REACTIVE FELINE PROTEIN
TITLE OF INVENTION: 103
NUMBER OF SEQUENCES: 103
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD, LLP
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/431,184
FILING DATE: 28-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/430,014
FILING DATE: 27-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/300,928
FILING DATE: 02-SEP-1994
ATTORNEY/AGENT INFORMATION:
NAME: Amy E. Mandragouras
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: IMI-044DV3

TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)742-4214
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 465 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..291
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 58..291
US-08-431-184-9

Query Match 56.0%; Score 16.8; DB 3; Length 465;
Best Local Similarity 75.0%; Pred. No. 47;
Matches 21; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 3 TGGCATGGGTAGAACGGCTGATACCCA 30
||| | | | | | | | | | | | | | |
Db 199 TGCTACGTGAGAACGCGACTCATATCCA 226

RESULT 10
US-08-300-928C-7
Sequence 7, Application US/08300928C
Patent No. 6019972
GENERAL INFORMATION:
APPLICANT: GERTER, Malcolm L. et al.
TITLE OF INVENTION: PEPTIDES FOR HUMAN T CELL REACTIVE FELINE
TITLE OF INVENTION: PROTEIN (TRFP)
NUMBER OF SEQUENCES: 101
CORRESPONDENCE ADDRESS:
ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION
STREET: 610 LINCOLN STREET
CITY: WALTHAM
STATE: Massachusetts
COUNTRY: USA
ZIP: 02145
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/300,928C
FILING DATE: September 2, 1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/807,529
FILING DATE: December 13, 1991
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: AMY E. MANDRAGOURAS
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: 002.60S(IMI-044)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 469 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..327

US-08-300-928C-7

Query Match

Best Local Similarity 56.0%; Score 16.8; DB 3; Length 469;
Matches 21; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 3 TGCATGGGTGAGACGGCGCTGATATCCA 30

Db 199 TGCTACGTGAGAACGGACGCTCATATCCA 226

RESULT 11

US-08-430-944D-7
; Sequence 7, Application US/08430944D
; Patent No. 6025162

; GENERAL INFORMATION:

; APPLICANT: Bruce L. Rogers et al.

; TITLE OF INVENTION: A HUMAN T CELL REACTIVE FELINE PROTEIN

; NUMBER OF SEQUENCES: 103

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: LAHIVE & COCKFIELD, LLP

; STREET: 28 State Street

; CITY: Boston

; STATE: Massachusetts

; COUNTRY: USA

; ZIP: 02109

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent in Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/430,944D

; FILING DATE: 28-APR-1995

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER: US 08/430,014

; FILING DATE: 27-APR-1995

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/300,928

; FILING DATE: 02-SEPT-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Amy E. Mandragouras

; REGISTRATION NUMBER: 36,207

; REFERENCE/DOCKET NUMBER: IMI-044DV2

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (617)227-7400

; TELEFAX: (617)742-4214

; INFORMATION FOR SEQ ID NO: 7:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 469 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 1..327

; NAME/KEY: mat_peptide

; LOCATION: 59..327

US-08-430-944D-7

Query Match

Best Local Similarity 75.0%; Score 16.8; DB 3; Length 469;
Matches 21; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

RESULT 12

;

US-08-430-014-7

; Sequence 7, Application US/08430014

; Patent No. 6048962

; GENERAL INFORMATION:

; APPLICANT: GEFTER, Malcolm L. et al.

; TITLE OF INVENTION: PEPTIDES FOR HUMAN T CELL REACTIVE FELINE

; NUMBER OF SEQUENCES: 101

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: IMMUNOLOGIC PHARMACEUTICAL CORPORATION

; STREET: 610 LINCOLN STREET

; CITY: WALTHAM

; STATE: Massachusetts

; COUNTRY: USA

; ZIP: 02145

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: ASCII text

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/430,014

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/300,928

; FILING DATE:

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: Amy E. Mandragouras

; REGISTRATION NUMBER: 36,207

; REFERENCE/DOCKET NUMBER: 002.6US(IMI-044)

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (617) 227-7400

; TELEFAX: (617) 227-5941

; INFORMATION FOR SEQ ID NO: 7:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 469 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 1..327

US-08-430-014-7

Query Match

Best Local Similarity 56.0%; Score 16.8; DB 3; Length 469;
Matches 21; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 3 TGCATGGGTGAGACGGCGCTGATATCCA 30

Db 199 TGCTACGTGAGAACGGACGCTCATATCCA 226

RESULT 13

US-08-431-184-7

; Sequence 7, Application US/08431184

; Patent No. 6120769

; GENERAL INFORMATION:

; APPLICANT: Bruce L. Rogers et al.

; TITLE OF INVENTION: A HUMAN T CELL REACTIVE FELINE PROTEIN

; NUMBER OF SEQUENCES: 103

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: LAHIVE & COCKFIELD, LLP

; STREET: 28 State Street

; CITY: Boston

; STATE: Massachusetts

; COUNTRY: USA

; ZIP: 02109

; COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/431,184
FILING DATE: 28-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/430,014
FILING DATE: 27-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/300,928
FILING DATE: 02-SEPT-1994
ATTORNEY/AGENT INFORMATION:
NAME: Amy E. Mandragouras
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: IMI-044DV3
TELEPHONE: (617)742-4214
TELEFAX: (617)227-7400
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 469 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..327
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 59..327
US-08-431-184-7

Query Match 56.0%; Score 16.8; DB 3; length 469;
Best Local Similarity 75.0%; Pred. No. 47;
Matches 21; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 3 TGCCATGGGTAGACGGCGCTGATACCA 30
||| | | | | | | | | | | | | | |
Db 199 TGCTACGTGAGAACGAGCTCATATCCA 226

RESULT 14
US-08-300-928C-5
Sequence 5, Application US/08300928C
Patent No. 6019972
GENERAL INFORMATION:
APPLICANT: GEFTER, Malcolm L. et al.
TITLE OF INVENTION: PEPTIDES FOR HUMAN T CELL REACTIVE FELINE
PROTEIN (TRFP)
NUMBER OF SEQUENCES: 101
CORRESPONDENCE ADDRESS:
ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION
STREET: 610 LINCOLN STREET
CITY: WALTHAM
STATE: Massachusetts
COUNTRY: USA
ZIP: 02145
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/300,928C
FILING DATE: September 2, 1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/807,529
FILING DATE: December 13, 1991
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: AMY E. MANDRAGOURAS
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: 002.605(IMI-044)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 476 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..335
US-08-300-928C-5

Query Match 56.0%; Score 16.8; DB 3; length 476;
Best Local Similarity 75.0%; Pred. No. 48;
Matches 21; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 3 TGCCATGGGTAGACGGCGCTGATACCA 30
||| | | | | | | | | | | | | | |
Db 200 TGCTACGTGAGAACGAGCTCATATCCA 227

RESULT 15
US-08-430-944D-5
Sequence 5, Application US/08430944D
Patent No. 6025162
GENERAL INFORMATION:
APPLICANT: Bruce L. Rogers et al.
TITLE OF INVENTION: A HUMAN T CELL REACTIVE FELINE PROTEIN
NUMBER OF SEQUENCES: 103
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD, LLP
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/430,944D
FILING DATE: 28-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/430,014
FILING DATE: 27-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/300,928
FILING DATE: 02-SEPT-1994
ATTORNEY/AGENT INFORMATION:
NAME: Amy E. Mandragouras
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: IMI-044DV2
TELEPHONE: (617)227-7400
TELEFAX: (617)742-4214
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 476 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:

NAME/KEY: CDS
 LOCATION: 2..334
 FEATURE:
 NAME/KEY: mat_peptide
 LOCATION: 59..334
 US-08-430-944D-5

Query Match 56.0%; Score 16.8; DB 3; Length 476;
 Best Local Similarity 75.0%; Pred. No. 48;
 Matches 21; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 QY 3 TGGCATGGCTAGACGGGCTGATACCCA 30
 Db 200 TGCTACGTGGAGACGGGCTGATACCCA 227

Search completed: November 23, 2002, 18:44:54
 Job time : 57 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OK nucleic - nucleic search, using sw model

Run on: November 23, 2002, 18:12:07 ; Search time 50 Seconds
(without alignments)
227.238 Million cell updates/sec

Title: US-09-848-781-4

Perfect score: 30

Sequence: 1 catgcattggtagacagcggtgtacacca 30

Scoring table:

Gapop 10.0 , Gapext 1.0

Searched: 335578 seqs, 18936513 residues

Total number of hits satisfying chosen parameters: 671156

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Published_Applications_NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCr_NEM_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCr_NEM_PUB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18.4	61.3	570	US-09-864-761-16234	Sequence 16234, A
2	18	60.0	143306	US-09-729-920-3	Sequence 3, Appli
3	18	60.0	465237	US-09-933-267a-1	Sequence 1, Appli
4	17.8	59.3	321	US-09-880-107-688	Sequence 888, App
5	17.8	59.3	615	US-10-037-860-6	Sequence 6, Appli
6	17.8	59.3	936	US-09-815-242-7776	Sequence 7776, Ap
7	17.8	59.3	1574	US-10-037-860-10	Sequence 10, Appl
8	17.8	59.3	3711	US-09-917-800A-1698	Sequence 1698, Ap
9	17.8	59.3	245487	US-10-026-188-3	Sequence 3, Appli
10	17.2	57.3	344	US-09-867-550-1729	Sequence 1729, Ap
11	17.2	57.3	615	US-09-893-737-213	Sequence 213, App
12	17.2	57.3	615	US-09-893-737-233	Sequence 233, App
13	17.2	57.3	627	US-10-104-611-31	Sequence 31, Appl
14	17.2	57.3	627	US-10-112-547-31	Sequence 31, Appl
15	17.2	57.3	627	US-10-112-241-31	Sequence 31, Appl
16	17.2	57.3	962	US-09-864-761-1805	Sequence 1805, Ap
17	17	56.7	403	US-09-960-253-113	Sequence 113, App
18	17	56.7	582	US-09-864-761-9318	Sequence 9318, Ap
19	17	56.7	21784	US-09-820-002-3	Sequence 3, Appli

20	16.8	56.0	467	10	US-09-867-701-10365	Sequence 10365, A
21	16.8	56.0	3090	9	US-09-938-842A-269	Sequence 269, App
22	16.8	56.0	22452	9	US-09-764-868-1487	Sequence 1487, Ap
23	16.8	56.0	22452	9	US-09-764-868-1489	Sequence 1489, Ap
24	16.6	55.3	448	10	US-09-867-701-10183	Sequence 10183, A
25	16.6	55.3	2784	10	US-09-954-456-772	Sequence 272, App
26	16.6	55.3	26668	10	US-09-962-832-222	Sequence 222, App
27	16.4	54.7	85	12	US-10-024-997-10	Sequence 10, Appl
28	16.4	54.7	86	12	US-10-024-997-11	Sequence 11, Appl
29	16.4	54.7	420	10	US-09-815-343-893	Sequence 893, App
30	16.4	54.7	422	10	US-09-925-299-203	Sequence 203, App
31	16.4	54.7	471	10	US-09-770-444-175	Sequence 175, App
32	16.4	54.7	500	10	US-09-864-761-2551	Sequence 2551, Ap
33	16.4	54.7	583	10	US-09-815-343-1175	Sequence 1175, Ap
34	16.4	54.7	2000	9	US-09-938-842A-3745	Sequence 3745, Ap
35	16.4	54.7	2090	10	US-09-764-847-1842	Sequence 1842, Ap
36	16.4	54.7	2091	10	US-09-764-847-1843	Sequence 1843, Ap
37	16.4	54.7	3112	8	US-08-927-939-34	Sequence 34, Appl
38	16.4	54.7	3839	10	US-09-815-656-57	Sequence 57, Appl
39	16.4	54.7	5140	12	US-10-044-090-94	Sequence 94, Appl
40	16.4	54.7	6902	10	US-09-764-847-1019	Sequence 1019, Appl
41	16.4	54.7	22161	10	US-09-764-847-1020	Sequence 1020, Ap
42	16.2	54.0	105	10	US-09-783-590-8073	Sequence 8073, Ap
43	16.2	54.0	291	10	US-09-764-869-53	Sequence 53, Appl
44	16.2	54.0	387	9	US-09-933-797-176	Sequence 176, App
45	16.2	54.0	401	9	US-09-946-807-792	Sequence 792, App

ALIGNMENTS

RESULT 1

US-09-864-761-16234

Sequence 16234, Application US/09864761

Patent No. US20020048763A1

GENERAL INFORMATION:

APPLICANT: Penn, Sharon G.

APPLICANT: Rank, David R.

APPLICANT: Hanzel, David K.

APPLICANT: Chen, Wensheng

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY

FILE REFERENCE: Aeomic-X-1

CURRENT APPLICATION NUMBER: US/09/864,761

CURRENT FILING DATE: 2001-05-23

PRIOR APPLICATION NUMBER: US 60/180,312

PRIOR FILING DATE: 2000-02-04

PRIOR APPLICATION NUMBER: US 60/207,456

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: US 09/632,366

PRIOR FILING DATE: 2000-08-03

PRIOR APPLICATION NUMBER: GB 24263.6

PRIOR FILING DATE: 2000-10-04

PRIOR APPLICATION NUMBER: US 60/236,359

PRIOR FILING DATE: 2000-09-27

PRIOR APPLICATION NUMBER: PCT/US01/00666

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00667

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00664

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00669

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00665

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00668

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00663

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00662

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00661

PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 4917
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 16234
; LENGTH: 570
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AP00011.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.6
US-09-864-761-16234

Query Match 61.3%; Score 18.4; DB 10; Length 570;
Best Local Similarity 78.6%; Pred. No. 9.1;
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 2 ATGCCATGGTAGACGGCTGATACCC 29
Db 433 ATGCCATGGGAGAACCTGCTCAATCC 460

RESULT 2
US-09-729-920-3
; Sequence 3, Application US/09729920
; Patent No. US20020103115A1
; GENERAL INFORMATION:
; APPLICANT: GUEGLER, Karl et al
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL000858
; CURRENT APPLICATION NUMBER: US/09/729,920
; CURRENT FILING DATE: 2000-12-06
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 143306
; TYPE: DNA
; ORGANISM: Human
US-09-729-920-3

Query Match 60.0%; Score 18; DB 10; Length 143306;
Best Local Similarity 80.8%; Pred. No. 42;
Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 5 CCATGGTAGACGGCTGATACCA 30
Db 25643 CAAGGATAGAAAGGCTGAGACCA 25668

RESULT 3
US-09-933-267A-1
; Sequence 1, Application US/09933267A
; Patent No. US20020123095A1
; GENERAL INFORMATION:
; APPLICANT: Kalush, Francis et al.
; TITLE OF INVENTION: Estrogen receptor alpha variants and
; TITLE OF INVENTION: Methods of detection thereof
; FILE REFERENCE: CL000258C14
; CURRENT APPLICATION NUMBER: US/09/933,267A
; CURRENT FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: 60/160626
; PRIOR FILING DATE: 1999-10-20
; PRIOR APPLICATION NUMBER: 60/183756
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: 09/692414

; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/768184
; PRIOR FILING DATE: 2001-01-24
; PRIOR APPLICATION NUMBER: 09/804076
; PRIOR FILING DATE: 2001-03-13
; PRIOR APPLICATION NUMBER: 09/826314
; PRIOR FILING DATE: 2001-04-05
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 465237
; TYPE: DNA
; ORGANISM: human
US-09-933-267A-1

Query Match 60.0%; Score 18; DB 10; Length 465237;
Best Local Similarity 80.8%; Pred. No. 52;
Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 3 TGCCATGGTAGACGGCTGATACC 28
Db 224733 TGCCATGGTAATGGGCTGTACC 224758

RESULT 4
US-09-880-107-688
; Sequence 688, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherif, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 688
; LENGTH: 321
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 AA287393
US-09-880-107-688

Query Match 59.3%; Score 17.8; DB 10; Length 321;
Best Local Similarity 75.9%; Pred. No. 16;
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 2 ATGCCATGGTAGACGGCTGATACCA 30
Db 280 AGGCCATGGGAGAGAGCCCTGATACCA 308

RESULT 5
US-10-037-860-6/C
; Sequence 6, Application US/10037860
; Patent No. US20020123114A1
; GENERAL INFORMATION:
; APPLICANT: Jerome B. Posner
; APPLICANT: Josep O. Dalmau
; APPLICANT: Myrna R. Rosenfeld
; TITLE OF INVENTION: MA FAMILY POLYPEPTIDES AND ANTI-MA
; TITLE OF INVENTION: ANTIBODIES
; FILE REFERENCE: 2581.1004-004
; CURRENT APPLICATION NUMBER: US/10/037,860
; CURRENT FILING DATE: 2001-01-04

```

; PRIOR APPLICATION NUMBER: 09/189,527
; PRIOR FILING DATE: 1998-11-10
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 615
; TYPE: DNA
; ORGANISM: homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(585)
US-10-037-860-6

```

```

Query Match          59.3%; Score 17.8; DB 12; Length 615;
Best Local Similarity 75.9%; Pred. No. 18;
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```

```

QY      2 ATGCCATGGGTAGAACGGGCTGATACCCA 30
Db      481 ATCTCATGGGTAGACGAGGGCTGAGCGCCA 453

```

```

RESULT 6
; Sequence 7776, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7776
; LENGTH: 936
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(936)
US-09-815-242-7776

```

```

Query Match          59.3%; Score 17.8; DB 10; Length 936;
Best Local Similarity 75.9%; Pred. No. 19;
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY      1 CATGCCATGGGTAGAACGGGCTGATACCC 29
Db      739 CCTGCCAGAGGTTGAACGGGGGTATCCAC 711

```

```

RESULT 7
; Sequence 10, Application US/10037860
; Patent No. US20020123114A1
; GENERAL INFORMATION:
; APPLICANT: Jerome B. Posner
; APPLICANT: Joseph O. Dalman
; APPLICANT: Myrna R. Rosenfield
; TITLE OF INVENTION: MA FAMILY POLYPEPTIDES AND ANTI-MA
; FILE REFERENCE: 2581.1004-004
; CURRENT APPLICATION NUMBER: US/10/037,860
; CURRENT FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 09/189,527
; PRIOR FILING DATE: 1998-11-10
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 1574
; TYPE: DNA
; ORGANISM: homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)...(850)
US-10-037-860-10

```

```

Query Match          59.3%; Score 17.8; DB 12; Length 1574;
Best Local Similarity 75.9%; Pred. No. 21;
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY      2 ATGCCATGGGTAGAACGGGCTGATACCCA 30
Db      242 ATCTCATGGGTAGACGAGGGCTGAGCGCCA 214

```

```

RESULT 8
; Sequence 1698, Application US/09917800A
; Patent No. US20020119462A1
; GENERAL INFORMATION:
; APPLICANT: Mendrick, Donna
; APPLICANT: Porter, Mark
; APPLICANT: Johnson, Kory
; APPLICANT: Castile, Arthur
; APPLICANT: Elashoff, Michael
; TITLE OF INVENTION: Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5038-US
; CURRENT APPLICATION NUMBER: US/09/917,800A
; CURRENT FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 60/222,040
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: US 60/222,880
; PRIOR FILING DATE: 2000-11-02
; PRIOR APPLICATION NUMBER: US 60/290,029
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/290,645
; PRIOR FILING DATE: 2001-05-15
; PRIOR APPLICATION NUMBER: US 60/292,336
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/295,798
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US 60/297,457
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,884
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/303,459
; PRIOR FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 1740
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 1698
; LENGTH: 3711

```

```
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020119462A1 NM_022287
US-09-917-800A-1698

Query Match
Best Local Similarity 59.3%; Score 17.8; DB 10; Length 3711;
Best Local Similarity 75.9%; Pred. No. 25;
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 ATGCATGGGTAGAACGGCTGATACCA 30
Db 2198 ATGCAGAGGTACACACGACCATCCCA 2170

RESULT 9
US-10-026-188-3/c
; Sequence 3, Application US/10026188
; Patent No. US20020164645A1
; GENERAL INFORMATION:
; APPLICANT: Zuker, Charles S.
; APPLICANT: Zhang, Yifeng
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Assays for Taste Receptor Cell Specific
; TITLE OF INVENTION: Ion Channel
; FILE REFERENCE: 02307E-114910US
; CURRENT APPLICATION NUMBER: US/10/026,188
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: US 60/259,379
; PRIOR FILING DATE: 2000-12-29
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 249487
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; OTHER INFORMATION: mouse genomic region containing 1trpc5
US-10-026-188-3

Query Match
Best Local Similarity 59.3%; Score 17.8; DB 9; Length 249487;
Best Local Similarity 75.9%; Pred. No. 58;
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CATGCCATGGGTAGAACGGCTGATACCC 29
Db 179708 CATGCTGTGATAGAAAGCAGATACTC 179680

RESULT 10
US-09-867-550-1729/c
; Sequence 1729, Application US/09867550
; Patent No. US20020082206A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Mehraban, Fuad,
; APPLICANT: Conley, Pamela
; APPLICANT: Law, Debbie
; APPLICANT: Topper, James
; TITLE OF INVENTION: No. US00020082206A1el Polynucleotides from Atherogenic Cells and
; TITLE OF INVENTION: Thereby
; FILE REFERENCE: 21402-013 (Cura-313)
; CURRENT APPLICATION NUMBER: US/09/867,550
; CURRENT FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: USSN 60/208,427
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 2125
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1729
; LENGTH: 344
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
```

```
; NAME/KEY: misc_feature
; LOCATION: (344)
; OTHER INFORMATION: wherein n is one of a or t or c or g
US-09-867-550-1729

Query Match
Best Local Similarity 57.3%; Score 17.2; DB 10; Length 344;
Best Local Similarity 73.3%; Pred. No. 31;
Matches 22; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CATGCCATGGGTAGAACGGCTGATACCA 30
Db 326 CATCCCATGGGACGACGCCCGGATCCA 297

RESULT 11
US-09-893-737-213/c
; Sequence 213, Application US/09893737
; Patent No. US20020110855A1
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Presnell, Scott R.
; TITLE OF INVENTION: MAMMALIAN SECRETED PROTEINS
; FILE REFERENCE: 00-41
; CURRENT APPLICATION NUMBER: US/09/893,737
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: US 60/215,446
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 329
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 213
; LENGTH: 615
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(615)
US-09-893-737-213

Query Match
Best Local Similarity 57.3%; Score 17.2; DB 10; Length 615;
Best Local Similarity 73.3%; Pred. No. 34;
Matches 22; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CATGCCATGGGTAGAACGGCTGATACCA 30
Db 445 CATGCCCTGATGTAATGTCTGGACAA 416

RESULT 12
US-09-893-737-233/c
; Sequence 233, Application US/09893737
; Patent No. US20020110855A1
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Presnell, Scott R.
; TITLE OF INVENTION: MAMMALIAN SECRETED PROTEINS
; FILE REFERENCE: 00-41
; CURRENT APPLICATION NUMBER: US/09/893,737
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: US 60/215,446
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 329
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 233
; LENGTH: 615
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(615)
US-09-893-737-233

Query Match
Best Local Similarity 57.3%; Score 17.2; DB 10; Length 615;
Best Local Similarity 73.3%; Pred. No. 34;
```

Matches 22: Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 CATGCCATGGTAGAAGCGGCTGATACCA 30
||||| | | | | | | | | | | | | | | |
Db 445 CATGCCCTAGGTGTAATGTCTGGACCAA 416

RESULT 13
US-10-104-611-31/c

; Sequence 31, Application US/10104611
; Patent No. US20020160976A1

GENERAL INFORMATION:

APPLICANT: Miles, Vincent J.
Mathews, Michael B.

Katze, Michael G.

Witherell, Gary

Watson, Julia C.

TITLE OF INVENTION: METHOD FOR SELECTIVE INACTIVATION
OF VIRAL REPLICATION

NUMBER OF SEQUENCES: 33

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Pennie & Edmonds

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10036/2711

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/104.611

FILING DATE: 22-Mar-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/221.816B

FILING DATE: 01-APR-1994

ATTORNEY/AGENT INFORMATION:

NAME: Coruzzi, Laura A

REGISTRATION NUMBER: 30,742

REFERENCE/DOCKET NUMBER: 7960-030

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 790-9090

TELEFAX: (212) 869-8864

TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 31:

SEQUENCE CHARACTERISTICS:

LENGTH: 627 base pairs

TYPE: nucleic acid

STRANDEDNESS: Single

TOPOLOGY: Linear

MOLECULE TYPE: RNA

SEQUENCE DESCRIPTION: SEQ ID NO: 31:

US-10-104-611-31

Query Match 57.3%; Score 17.2; DB 9; Length 627;
Best Local Similarity 73.3%; Pred. No. 35;
Matches 22: Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 CATGCCATGGTAGAAGCGGCTGATACCA 30
|| | | | | | | | | | | | | | | |
Db 43 CACCCATGGTGTGATGTGGATACCA 14

RESULT 14
US-10-112-547-31/c

; Sequence 31, Application US/10112547
; Patent No. US20020160977A1

GENERAL INFORMATION:

APPLICANT: Miles, Vincent J.
Mathews, Michael B.

Katze, Michael G.

Witherell, Gary

Watson, Julia C.

TITLE OF INVENTION: METHOD FOR SELECTIVE INACTIVATION
OF VIRAL REPLICATION

NUMBER OF SEQUENCES: 33

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Pennie & Edmonds

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10036/2711

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/112.547

FILING DATE: 28-Mar-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/221.816B

FILING DATE: 01-APR-1994

ATTORNEY/AGENT INFORMATION:

NAME: Coruzzi, Laura A

REGISTRATION NUMBER: 30,742

REFERENCE/DOCKET NUMBER: 7960-030

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 790-9090

TELEFAX: (212) 869-8864

TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 31:

SEQUENCE CHARACTERISTICS:

LENGTH: 627 base pairs

TYPE: nucleic acid

STRANDEDNESS: Single

TOPOLOGY: Linear

MOLECULE TYPE: RNA

SEQUENCE DESCRIPTION: SEQ ID NO: 31:

US-10-112-547-31

Query Match 57.3%; Score 17.2; DB 9; Length 627;
Best Local Similarity 73.3%; Pred. No. 35;
Matches 22: Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 CATGCCATGGTAGAAGCGGCTGATACCA 30
|| | | | | | | | | | | | | | | |
Db 43 CACCCATGGTGTGATGTGGATACCA 14

RESULT 15
US-10-112-241-31/c

; Sequence 31, Application US/10112241
; Patent No. US20020165194A1

GENERAL INFORMATION:

APPLICANT: Miles, Vincent J.
Mathews, Michael B.

Katze, Michael G.

Witherell, Gary

Watson, Julia C.

TITLE OF INVENTION: METHOD FOR SELECTIVE INACTIVATION
OF VIRAL REPLICATION

NUMBER OF SEQUENCES: 33

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Pennie & Edmonds

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10036/2711

COMPUTER READABLE FORM:

```

; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/112, 241
; FILING DATE: 28-Mar-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/221, 816B
; FILING DATE: 01-Apr-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7960-030
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 627 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: RNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 31:
US-10-112-241-31

```

```

Query Match      57.3%; Score 17.2; DB 9; Length 627;
Best Local Similarity 73.3%; Pred. No. 35;
Matches 22; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
OY      1 CATGCCATGGTGAACGGGCTGATACCA 30
      11 | | | | | | | | | | | | | | | |
Db      43 CACCCAATGGGTGATGTTGGGATACCA 14

```

Search completed: November 23, 2002, 19:44:04
Job time : 209 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 17:20:01 ; Search time 2002 Seconds
(without alignments)
242.690 Million cell updates/sec

Title: US-09-848-781-4
Perfect score: 30
Sequence: 1 catgcacggtgtagaacggctgataccca 30

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 segs, 8097743376 residues
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	66.7	782	14	BQ955349 AGENCOURT
2	19.6	65.3	215	17	BH321569 CH230-790
3	19.6	65.3	547	17	AQ406947 HS_5113_A
4	19.4	64.7	349	17	AQ059495 CIT-HSP-2
5	19.4	64.7	420	17	AQ005523 CIT-HSP-2
6	19.4	64.7	435	10	BE378910 601237490

	7	19.4	64.7	669	13	BI147497
	8	19.4	64.7	711	17	A2701481
	9	19	63.3	435	10	AM634546
	10	19	63.3	560	10	AM638163
	11	19	63.3	846	10	AV751935
	12	18.8	62.7	199	14	793335
	13	18.8	62.7	312	9	A1046979
	14	18.8	62.7	322	12	BF088838
	15	18.8	62.7	323	9	AL367239
	16	18.8	62.7	403	9	AA114706
	17	18.8	62.7	403	9	AL384668
	18	18.8	62.7	410	9	AA266270
	19	18.8	62.7	434	9	A1322907
	20	18.8	62.7	447	9	AL366759
	21	18.8	62.7	478	10	AM825699
	22	18.8	62.7	486	9	AA170581
	23	18.8	62.7	501	9	AA839657
	24	18.8	62.7	501	9	AL387211
	25	18.8	62.7	506	12	BF022057
	26	18.8	62.7	507	9	AA553364
	27	18.8	62.7	515	10	BE689144
	28	18.8	62.7	516	13	BI343249
	29	18.8	62.7	518	9	AL369105
	30	18.8	62.7	520	9	AA213089
	31	18.8	62.7	524	9	AL369693
	32	18.8	62.7	528	12	BE144714
	33	18.8	62.7	529	10	AM968610
	34	18.8	62.7	539	9	A1893613
	35	18.8	62.7	549	9	AL365824
	36	18.8	62.7	551	9	AL368464
	37	18.8	62.7	552	9	AL375820
	38	18.8	62.7	559	12	BF730095
	39	18.8	62.7	563	12	BE722382
	40	18.8	62.7	594	9	A1787444
	41	18.8	62.7	600	12	BE585296
	42	18.8	62.7	614	10	AM044919
	43	18.8	62.7	648	10	AM044936
	44	18.8	62.7	659	17	A2383179
	45	18.8	62.7	713	10	BE284742

ALIGNMENTS

RESULT 1
BQ955349
LOCUS
DEFINITION
AGENCOURT_8732971 NIH_MGC_101 Homo sapiens cDNA clone IMAGE:6455241
5', mRNA sequence.
ACCESSION
BQ955349
VERSION
BQ955349.1 GI:22370827
KEYWORDS
SOURCE
Homo sapiens
human.
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 782)
NIH-MGC <http://mgc.nci.nih.gov/>.
AUTHORS
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LMNL at:
<http://image.llnl.gov>
Plate: L10C2633 row: 1 column: 10
High quality sequence stop: 407.
Location/Qualifiers
1..782

FEATURES

source

```
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6455241"
/clone.lib="NIH_MGC.101"
/tissue_type="epidermoid carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: Lung; Vector: pOTB7; Site_1: EcoRI; Site_2:
XhoI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC library."
```

BASE COUNT 237 a 176 c 206 g 155 t 8 others

ORIGIN

Query Match 66.7%; Score 20; DB 14; Length 782;
Best Local Similarity 82.1%; Pred. No. 3.7e+02;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 3 TGGCATGGTGAACGGGCTGATACCA 30
||||| ||| | ||||| ||| |||
Db 623 TGCCAGCGGAAAAAGGCTGATATCCA 650

RESULT 2
BH321569 215 bp DNA linear GSS 03-DEC-2001
LOCUS CH230-79020.T1 CHORI-230 Segment 1 Rattus norvegicus genomic clone
DEFINITION CH230-79020, DNA sequence.
ACCESSION BH321569
VERSION BH321569.1 GI:17252283
KEYWORDS GSS.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 215)
Zhao,S., Shetty,J., Shatsman,S., Tsegaye,G., Geer,K., Shvartsbeyn
,A., Gebregeorgis,E., Overton,L., Russell,D., Chen,D., Riggs,F., de
Jong,P. and Fraser,C.M.
Rat BAC End Sequences from Library CHORI-230 EcoRI segment
Unpublished (1999)
Other_GSS: CH230-79020.TV
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tifgr.org
Clones are derived from the rat BAC library CHORI-230
(<http://www.chori.org/bacpac/rat230.htm>). For BAC library
availability, please contact Pieter de Jong (pdejong@mail.cbo.org).
Clones may be purchased from BACPAC Resources
(<http://www.chori.org/bacpac/orering/information.htm>). BAC end
plate: http://www.tifgr.org/tidb/bac_ends/rat/bac_end_intro.html
Plate: 79 row: 0 column: 20
Seq primer: SP6
Class: BAC ends.
Location/Qualifiers
1. 215
/organism="Rattus norvegicus"
/strain="BN/SSNhsd/MCW"
/db_xref="taxon:10116"
/clone="CH230-79020"
/clone.lib="CHORI-230 Segment 1"
/sex="Female"
/cell_type="Brain"
/note="Vector: pTARBAC2.1; Site_1: EcoRI; Site_2: EcoRI;
CHORI-230 Rat (BN/SSNhsd/MCW) BAC library produced by

BASE COUNT 43 a 72 c 34 g 66 t

ORIGIN

Query Match 65.3%; Score 19.6; DB 17; Length 215;
Best Local Similarity 84.6%; Pred. No. 3.2e+02;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 3 TGGCATGGTGAACGGGCTGATACC 28
||||| ||||| ||||| ||| |||
Db 125 TGCCGTAGTAGACCGGCTGATCCC 150

RESULT 3
AQ406947 547 bp DNA linear GSS 17-MAR-1999
LOCUS HS-5113_A1_E03_T7A RPCI-11 Human Male BAC Library Homo sapiens
DEFINITION genomic clone Plate-689 Col-5 Row-1, DNA sequence.
ACCESSION AQ406947
VERSION AQ406947.1 GI:4429569
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 547)
Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
Hood,L.
Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
99380589
CONTACT: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieterdejong.med.bufileo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.bufileo.edu/ordering_bac.htm)
or from Resear h Genetics (inf@resgen.com). BAC end Web Server:
<http://www.htsc.washington.edu>
Plate: 689 row: I column: 5
Seq primer: T7
Class: BAC ends
High quality sequence stop: 547.
Location/Qualifiers
1. 547
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="Plate-689 Col-5 Row-1"
/clone.lib="RPCI-11 Human Male BAC Library"
/sex="male"
/note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
Male blood DNA was isolated from one randomly chosen donor
and partially digested with a combination of EcoRI and
EcoRI Methylase. Size selected DNA was cloned into the
pBACe3.6 vector at EcoRI sites"

BASE COUNT 165 a 103 c 90 g 174 t 15 others

ORIGIN

Query Match 65.3%; Score 19.6; DB 17; Length 547;
Best Local Similarity 75.9%; Pred. No. 4.7e+02;
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 2 ATGCCATGGTGAACGGGCTGATACCA 30
||||| ||| | ||||| |||
Db 20 ATGCCATGGTGAACCCNGTGTATACNCA 48

RESULT 4										
LOCUS	A0059495	349 bp	DNA	linear	GSS 31-JUL-1998					
DEFINITION	CIT-HSP-2355L21.TR CIT-HSP Homo sapiens genomic clone 2355L21, DNA sequence.									
ACCESSION	A0059495									
VERSION	A0059495.1	GI:3361832								
KEYWORDS	GSS.									
SOURCE	human.									
ORGANISM	Homo sapiens									
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.									
AUTHORS	Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K., Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and Venter,J.C. Use of a random BAC End Sequence Database for Sequence-Ready Map Building (1998)									
TITLE	Unpublished (1998)									
JOURNAL	Other_GSSs: CIT-HSP-2355L21.TF									
COMMENT	Contact: Mark Adams Department of Eukaryotic Genomics The Institute for Genomic Research 9712 Medical Center Dr., Rockville, MD 20850, USA Tel: 301 838 0200 Fax: 301 838 0208 Email: mdadams@tigr.org Clones are available from Research Genetics (info@resgen.com). BAC end search page: http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html . Seq primer: M13 Reverse Class: BAC ends.									
FEATURES	Location/Qualifiers									
source	1..349									
	/organism="Homo sapiens"									
	/db_xref="taxon:9606"									
	/clone="2355L21"									
	/clone_lib="CIT-HSP"									
	/sex="Male"									
	/cell_type="Sperm"									
	/note="Vector: pbelOBAC11; site_1: HindIII; site_2: HindIII"									
BASE COUNT	98 a 72 c 74 g 104 t 1 others									
ORIGIN										
Query Match	64.7%; Score 19.4; DB 17; Length 349;									
Best Local Similarity	79.3%; Pred. NO. 4.7e+02;									
Matches 23; Conservative	0; Mismatches 6; Indels 0; Gaps 0;									
OY	2 ATGCCATGGGTGAACGGCGTGAATCCCA 30									
Db	114 ATGTATGTGGTATGAAGGATGCTACCCA 142									
RESULT 5										
LOCUS	A0005523	420 bp	DNA	linear	GSS 27-JUN-1998					
DEFINITION	CIT-HSP-2296E5.TR CIT-HSP Homo sapiens genomic clone 2296E5, DNA sequence.									
ACCESSION	A0005523									
VERSION	A0005523.1	GI:3082968								
KEYWORDS	GSS.									
SOURCE	human.									
ORGANISM	Homo sapiens									
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.									
AUTHORS	Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K., Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and Venter,J.C. Use of a random BAC End Sequence Database for Sequence-Ready Map Building (1998)									
TITLE										

```

JOURNAL      Unpublished (1998)
COMMENT      Other GSSS: CIT-HSP-2296E5.TF

Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Names are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
Seq primer: M13 Reverse
Class: BAC ends.

FEATURES
  source
    1..420
      /organism="Homo sapiens"
      /db_xref="GDB:7152897"
      /db_xref="taxon:9606"
      /clone="2296E5"
      /clone_1id="CIT-HSP"
      /sex="male"
      /cell_type="Sperm"
      /note="Vector: pBeloBAC11; site_1: HindIII; site_2:
HindIII"

BASE COUNT   122 a      85 c      87 g      126 t

ORIGIN
Query Match      64.7%; Score 19.4; DB 17; Length 420;
Best Local Similarity 79.3%; Pred. No. 5.1e+02;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Oy      2 ATGCCATGGGTAGAACGGCGCTGATACCA 30
      111 111111111 111 1111111
Db      150 ATGTAATGGGTAGTAGAAGATCTACCA 178
      mRNA sequence.

BE378910      435 bp      mRNA      linear      EST 21-JUL-2000
DEFINITION    601237490r1 NIH_MGC_44 Homo sapiens cDNA clone IMAGE:3609550 5',
LOCUS         BE378910
ACCESSION     BE378910
VERSION       BE378910.1 GI:9324275
KEYWORDS      EST.
SOURCE        human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 435)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strusberg, Ph.D.
Email: cga@bbs.femail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
CNHA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
http://image.lnl.gov
Plate: L1CM264 row: c column: 23
High quality sequence stop: 166.

FEATURES
  source
    1..435
      /organism="Homo sapiens"
      /db_xref="taxon:9606"
      /clone="IMAGE:3609550"
      /clone_1id="NIH_MGC_44"
      /tissue_type="endometrium, adenocarcinoma cell line"
      /lab_host="DH10B (phage-resistant)"
      /note="Organ: uterus; Vector: pOTB7; Site_1: XhoI; Site_2:

```


AUTHORS
 Blackshear, P.J., Lai, W.S., Thorn, J.M., Kennington, E.A., Staffa, N.G., Jr., Moore, D.T., Bouffard, G.G., Beckstrom-Sternberg, S.M., Touchman, J.W., Bonaldo, M.F., and Soares, M.B.
 The NIEHS Xenopus maternal EST project: interim analysis of the first 13,879 ESTs from unfertilized eggs

JOURNAL MEDLINE
 Gene 267 (1), 71-87 (2001)
 21211403

COMMENT
 Contact: Perry J. Blackshear
 Office of Clinical Research and Laboratory of Signal Transduction
 National Institute of Environmental Health Sciences
 A2-05 NIEHS, 101 Alexander Drive, Research Triangle Park, NC 27709, USA
 Tel: 919 541-4899
 Fax: 919 541-4571
 Email: black009@niehs.nih.gov
 Clone is available through Research Genetics, Inc., 2130 Memorial Parkway, Huntsville, AL 35901
 phone 800-533-4363 ext.cdna, fax 256-536-9016 att:cdna, email cdna@resgen.com
 DNA Sequencing and analyses performed by National Institutes of Health Intramural Sequencing Center (NISC).
 PCR Primers
 FORWARD: TGTAAACGACGCCAGT
 BACKWARD: CAGGAACAGCTATGACC
 Plate: 0022 row: A column: 06
 Seq primer: T7 primer.

FEATURES
 source
 Location/Qualifiers
 1..435
 /organism="Xenopus laevis"
 /db_xref="taxon:8355"
 /clone="PBX0022A06"
 /clone_lib="Blackshear/Soares normalized Xenopus egg library"
 /sex="female"
 /tissue_type="unfertilized egg"
 /cell_type="unfertilized egg"
 /dev_stage="unfertilized egg"
 /lab_host="DH10B"
 /note="Vector: pT7T3-Pac; Site_1: EcoRI; Site_2: NotI; PolyA-selected mRNA was prepared from unfertilized Xenopus laevis eggs. The library was constructed in the vector pT7T3-Pac as described in Bonaldo, M.F., Lennon, G. and Soares, M.B. 'Normalization and subtraction: two approaches to facilitate gene discovery', Genome Research 6:791-806, 1996. The first strand synthesis used a NotI-drf18 primer; double stranded cDNAs were ligated to EcoRI adapters, digested with NotI, and directionally cloned into the NotI and EcoRI-digested pT7T3-Pac vector. The library contained approximately 7.2 x 10⁵ recombinants, with average insert sizes of 1-1.5 kb."

BASE COUNT
 133 a 69 c 97 g 135 t 1 others

ORIGIN
 Query Match 63.3%; Score 19; DB 10; Length 435;
 Best Local Similarity 81.5%; Pred. No. 7.7e+02;
 Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY
 2 ATGCCATGGGTAGAACGGCGTGATACC 28
 ||||||| ||||| ||||| ||||| |||||

Db
 46 ATACCATGGCGACAGACACTAATACC 72

RESULT 10
 AM638163 560 bp mRNA linear EST 26-APR-2001
 LOCUS AM638163
 DEFINITION laevis cdna clone PBX006H03 5', mRNA sequence.
 ACCESSION AM638163
 VERSION AM638163.1 GI:7395304
 KEYWORDS EST.
 SOURCE African clawed frog.
 ORGANISM Xenopus laevis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
 1 (bases 1 to 560)
 Blackshear, P.J., Lai, W.S., Thorn, J.M., Kennington, E.A., Staffa, N.G., Jr., Moore, D.T., Bouffard, G.G., Beckstrom-Sternberg, S.M., Touchman, J.W., Bonaldo, M.F., and Soares, M.B.
 The NIEHS Xenopus maternal EST project: interim analysis of the first 13,879 ESTs from unfertilized eggs

JOURNAL MEDLINE
 Gene 267 (1), 71-87 (2001)
 21211403

COMMENT
 Contact: Perry J. Blackshear
 Office of Clinical Research and Laboratory of Signal Transduction
 National Institute of Environmental Health Sciences
 A2-05 NIEHS, 101 Alexander Drive, Research Triangle Park, NC 27709, USA
 Tel: 919 541-4899
 Fax: 919 541-4571
 Email: black009@niehs.nih.gov
 Clone is available through Research Genetics, Inc., 2130 Memorial Parkway, Huntsville, AL 35901
 phone 800-533-4363 ext.cdna, fax 256-536-9016 att:cdna, email cdna@resgen.com
 DNA Sequencing and analyses performed by National Institutes of Health Intramural Sequencing Center (NISC).
 PCR Primers
 FORWARD: TGTAAACGACGCCAGT
 BACKWARD: CAGGAACAGCTATGACC
 Plate: 0066 row: H column: 03
 Seq primer: T7 primer.

FEATURES
 source
 Location/Qualifiers
 1..560
 /organism="Xenopus laevis"
 /db_xref="taxon:8355"
 /clone="PBX006H03"
 /clone_lib="Blackshear/Soares normalized Xenopus egg library"
 /sex="female"
 /tissue_type="unfertilized egg"
 /cell_type="unfertilized egg"
 /dev_stage="unfertilized egg"
 /lab_host="DH10B"
 /note="Vector: pT7T3-Pac; Site_1: EcoRI; Site_2: NotI; PolyA-selected mRNA was prepared from unfertilized Xenopus laevis eggs. The library was constructed in the vector pT7T3-Pac as described in Bonaldo, M.F., Lennon, G. and Soares, M.B. 'Normalization and subtraction: two approaches to facilitate gene discovery', Genome Research 6:791-806, 1996. The first strand synthesis used a NotI-drf18 primer; double stranded cDNAs were ligated to EcoRI adapters, digested with NotI, and directionally cloned into the NotI and EcoRI-digested pT7T3-Pac vector. The library contained approximately 7.2 x 10⁵ recombinants, with average insert sizes of 1-1.5 kb."

BASE COUNT
 172 a 104 c 134 g 150 t

ORIGIN
 Query Match 63.3%; Score 19; DB 10; Length 560;
 Best Local Similarity 81.5%; Pred. No. 8.6e+02;
 Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY
 2 ATGCCATGGGTAGAACGGCGTGATACC 28
 ||||||| ||||| ||||| ||||| |||||

Db
 326 ATACCATGGCGACAGACACTAATACC 352

RESULT 11
 AV751935 846 bp mRNA linear EST 19-OCT-2000
 LOCUS AV751935
 DEFINITION NPD Homo sapiens CDNA clone NPDAXH10 5', mRNA sequence.
 ACCESSION AV751935
 VERSION AV751935.1 GI:10909783
 KEYWORDS EST.
 SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 846)
AUTHORS Song, H., Peng, Y., Gu, Y., Yang, Y., Gao, G., Xiao, H., Xu, X., Li, N., Qian, B., Liu, F., Qu, J., Gao, X., Cheng, Z., Xu, Z., Deng, L., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu, G., Ye, M., Zhang, Q., Han, Z., Chen, Z., Hu, R. and Chen, J.
Homo sapiens NPD library cDNA clones
Unpublished (2000)
COMMENT Contact: Qinhua Zhang
Shanghai Institute of Endocrinology, Rui-Jin Hospital
197 Rui-Jin II Road, Shanghai 200025, P. R. China
Tel: 86-21-64370045(ex.663332)
Fax: 86-21-64743206
Email: mshlms.stn.sh.cn
This clone is available at Shanghai Hematology Institute in Shanghai.
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong.
Location/Qualifiers
FEATURES
source 1..846
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="NPDAH10"
/clone_1ib="NPD"
/cissue_type="pituitary"
/dev_stage="Adult"
/lab_host="SOLR"
/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2: XhoI"
BASE COUNT 221 a 252 c 193 g 180 t
ORIGIN
Query Match 63.3%; Score 19; DB 10; Length 846;
Best Local Similarity 81.5%; Pred. No. 1e+03;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy 1 CATGCCATGGGTAGAACGGGCTGATAC 27
|||||
Db 720 CATGCCATGGGTACACAGCGCATAC 746
RESULT 12
T93355 199 bp mRNA linear EST 09-DEC-1998
LOCUS T93355 LMLV39cDNA Leishmania major cDNA clone Lm049 5' END, mRNA
DEFINITION
sequence.
ACCESSION T93355
VERSION T93355.1 GI:726528
KEYWORDS EST.
SOURCE Leishmania major
ORGANISM Leishmania major
Eukaryota; Euzlenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
REFERENCE 1 (bases 1 to 199)
AUTHORS Levick, M.P., Blackwell, J.M., Connor, V., Coulson, R.M., Miles, A., Smith, H.E., Wan, K.L. and Ajlaka, J.W.
An expressed sequence tag analysis of a full-length, spliced-leader cDNA library from Leishmania major promastigotes
Mol. Biochem. Parasitol. 76 (1-2), 345-348 (1996)
JOURNAL 97077459
COMMENT Contact: Blackwell JM
Cambridge Institute for Medical Research
Wellcome Trust/MRC Building, Addenbrooke's Hospital, Hills Road, Cambridge CB2 2X1, UK
Tel: 01223 336 143
Fax: 01223 331 206
Email: jmb37@cus.cam.ac.uk
Seq primer: T3.
FEATURES
source Location/Qualifiers
1..199
/organism="Leishmania major"

/strain="LV39"
/db_xref="taxon:5664"
/clone="Lm049"
/clone_1ib="LMLV39cDNA"
/lab_host="XLI-Blue MRF"
/note="Vector: Lambda UNIZAP; Site_1: NotI; Site_2: XhoI; Mid or late log promastigotes, full length; splice leader oligo used for second strand synthesis; directional cloning NotI at 5' end, XhoI at 3' end."
BASE COUNT 50 a 73 c 33 g 43 t
ORIGIN
Query Match 62.7%; Score 18.8; DB 14; Length 199;
Best Local Similarity 76.7%; Pred. No. 6.8e+02;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
Qy 1 CATGCCATGGGTAGAACGGGCTATACCA 30
|||||
Db 120 CATGGTATGGCGCACGCGCTGATACGA 91
RESULT 13
A1046979 312 bp mRNA linear EST 08-JUL-1998
LOCUS A1046979
DEFINITION
un53d02.r1 Soares-embryonic-stem-cell NMES Mus musculus cDNA clone IMAGE:1749123 5' similar to gb:X6503 ADENYLOSUCCLINATE SYNTHETASE (HUMAN); gb:M31654 Mouse growth hormone-releasing hormone mRNA, complete (MOUSE);, mRNA sequence.
ACCESSION A1046979
VERSION A1046979.1 GI:3295266
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 312)
AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.
The WashU-HM Mouse EST Project
Unpublished (1996)
Contact: Maria M/Mouse EST Project
WashU-HM Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LIND; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:960935
Seq primer: -28m3 rev2 ET from Amersham
High quality sequence stop: 311.
FEATURES
source Location/Qualifiers
1..312
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:1749123"
/clone_1ib="Soares-embryonic-stem-cell_NMES"
/cell_type="embryonic stem cell"
/lab_host="DH10B"
/note="Organ: bone marrow; Vector: pRT730-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTTACATCTGAGTGGAGCGCCCATGTGTTTTTTTTTTTTTTTTTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pRT730 vector. Library went through two rounds of normalization, and was constructed by Bento Soares and M.Fatima Bonaldo."

BASE COUNT 76 a 65 c 65 g 106 t
ORIGIN

Query Match 62.7%; Score 18.8; DB 9; Length 312;
Best Local Similarity 76.7%; Pred. No. 8.2e+02;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CATGCCATGGGTAGAACGGCGTGAATACCA 30
|||||
Db 72 CTGCGCATGGTTAGACACTGATATATCCCA 101

RESULT 14
BF088838
LOCUS 322 bp mRNA linear EST 19-OCT-2000
DEFINITION QY0-HN0900-140900-381-c04 HN0900 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF088838
VERSION BF088838.1 GI:10894548
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eulheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 322)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bata,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

JOURNAL MEDLINE
COMMENT 20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL:
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=6t2-QY0-HN0900-140
900-381-c04&t3=2000-09-14&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 37
High quality sequence stop: 267.
Location/Qualifiers
1..322
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_id="HN0900"
/dev_stage="Adult"
/note="Organ: head/neck; Vector: puc18; Site:1; Site:2; Site:3; Site:4; Site:5; Site:6; Site:7; Site:8; Site:9; Site:10; Site:11; Site:12; Site:13; Site:14; Site:15; Site:16; Site:17; Site:18; Site:19; Site:20; Site:21; Site:22; Site:23; Site:24; Site:25; Site:26; Site:27; Site:28; Site:29; Site:30; Site:31; Site:32; Site:33; Site:34; Site:35; Site:36; Site:37; Site:38; Site:39; Site:40; Site:41; Site:42; Site:43; Site:44; Site:45; Site:46; Site:47; Site:48; Site:49; Site:50; Site:51; Site:52; Site:53; Site:54; Site:55; Site:56; Site:57; Site:58; Site:59; Site:60; Site:61; Site:62; Site:63; Site:64; Site:65; Site:66; Site:67; Site:68; Site:69; Site:70; Site:71; Site:72; Site:73; Site:74; Site:75; Site:76; Site:77; Site:78; Site:79; Site:80; Site:81; Site:82; Site:83; Site:84; Site:85; Site:86; Site:87; Site:88; Site:89; Site:90; Site:91; Site:92; Site:93; Site:94; Site:95; Site:96; Site:97; Site:98; Site:99; Site:100; Site:101; Site:102; Site:103; Site:104; Site:105; Site:106; Site:107; Site:108; Site:109; Site:110; Site:111; Site:112; Site:113; Site:114; Site:115; Site:116; Site:117; Site:118; Site:119; Site:120; Site:121; Site:122; Site:123; Site:124; Site:125; Site:126; Site:127; Site:128; Site:129; Site:130; Site:131; Site:132; Site:133; Site:134; Site:135; Site:136; Site:137; Site:138; Site:139; Site:140; Site:141; Site:142; Site:143; Site:144; Site:145; Site:146; Site:147; Site:148; Site:149; Site:150; Site:151; Site:152; Site:153; Site:154; Site:155; Site:156; Site:157; Site:158; Site:159; Site:160; Site:161; Site:162; Site:163; Site:164; Site:165; Site:166; Site:167; Site:168; Site:169; Site:170; Site:171; Site:172; Site:173; Site:174; Site:175; Site:176; Site:177; Site:178; Site:179; Site:180; Site:181; Site:182; Site:183; Site:184; Site:185; Site:186; Site:187; Site:188; Site:189; Site:190; Site:191; Site:192; Site:193; Site:194; Site:195; Site:196; Site:197; Site:198; Site:199; Site:200; Site:201; Site:202; Site:203; Site:204; Site:205; Site:206; Site:207; Site:208; Site:209; Site:210; Site:211; Site:212; Site:213; Site:214; Site:215; Site:216; Site:217; Site:218; Site:219; Site:220; Site:221; Site:222; Site:223; Site:224; Site:225; Site:226; Site:227; Site:228; Site:229; Site:230; Site:231; Site:232; Site:233; Site:234; Site:235; Site:236; Site:237; Site:238; Site:239; Site:240; Site:241; Site:242; Site:243; Site:244; Site:245; Site:246; Site:247; Site:248; Site:249; Site:250; Site:251; Site:252; Site:253; Site:254; Site:255; Site:256; Site:257; Site:258; Site:259; Site:260; Site:261; Site:262; Site:263; Site:264; Site:265; Site:266; Site:267; Site:268; Site:269; Site:270; Site:271; Site:272; Site:273; Site:274; Site:275; Site:276; Site:277; Site:278; Site:279; Site:280; Site:281; Site:282; Site:283; Site:284; Site:285; Site:286; Site:287; Site:288; Site:289; Site:290; Site:291; Site:292; Site:293; Site:294; Site:295; Site:296; Site:297; Site:298; Site:299; Site:300; Site:301; Site:302; Site:303; Site:304; Site:305; Site:306; Site:307; Site:308; Site:309; Site:310; Site:311; Site:312; Site:313; Site:314; Site:315; Site:316; Site:317; Site:318; Site:319; Site:320; Site:321; Site:322; Site:323; Site:324; Site:325; Site:326; Site:327; Site:328; Site:329; Site:330; Site:331; Site:332; Site:333; Site:334; Site:335; Site:336; Site:337; Site:338; Site:339; Site:340; Site:341; Site:342; Site:343; Site:344; Site:345; Site:346; Site:347; Site:348; Site:349; Site:350; Site:351; Site:352; Site:353; Site:354; Site:355; Site:356; Site:357; Site:358; Site:359; Site:360; Site:361; Site:362; Site:363; Site:364; Site:365; Site:366; Site:367; Site:368; Site:369; Site:370; Site:371; Site:372; Site:373; Site:374; Site:375; Site:376; Site:377; Site:378; Site:379; Site:380; Site:381; Site:382; Site:383; Site:384; Site:385; Site:386; Site:387; Site:388; Site:389; Site:390; Site:391; Site:392; Site:393; Site:394; Site:395; Site:396; Site:397; Site:398; Site:399; Site:400; Site:401; Site:402; Site:403; Site:404; Site:405; Site:406; Site:407; Site:408; Site:409; Site:410; Site:411; Site:412; Site:413; Site:414; Site:415; Site:416; Site:417; Site:418; Site:419; Site:420; Site:421; Site:422; Site:423; Site:424; Site:425; Site:426; Site:427; Site:428; Site:429; Site:430; Site:431; Site:432; Site:433; Site:434; Site:435; Site:436; Site:437; Site:438; Site:439; Site:440; Site:441; Site:442; Site:443; Site:444; Site:445; Site:446; Site:447; Site:448; Site:449; Site:450; Site:451; Site:452; Site:453; Site:454; Site:455; Site:456; Site:457; Site:458; Site:459; Site:460; Site:461; Site:462; Site:463; Site:464; Site:465; Site:466; Site:467; Site:468; Site:469; Site:470; Site:471; Site:472; Site:473; Site:474; Site:475; Site:476; Site:477; Site:478; Site:479; Site:480; Site:481; Site:482; Site:483; Site:484; Site:485; Site:486; Site:487; Site:488; Site:489; Site:490; Site:491; Site:492; Site:493; Site:494; Site:495; Site:496; Site:497; Site:498; Site:499; Site:500; Site:501; Site:502; Site:503; Site:504; Site:505; Site:506; Site:507; Site:508; Site:509; Site:510; Site:511; Site:512; Site:513; Site:514; Site:515; Site:516; Site:517; Site:518; Site:519; Site:520; Site:521; Site:522; Site:523; Site:524; Site:525; Site:526; Site:527; Site:528; Site:529; Site:530; Site:531; Site:532; Site:533; Site:534; Site:535; Site:536; Site:537; Site:538; Site:539; Site:540; Site:541; Site:542; Site:543; Site:544; Site:545; Site:546; Site:547; Site:548; Site:549; Site:550; Site:551; Site:552; Site:553; Site:554; Site:555; Site:556; Site:557; Site:558; Site:559; Site:560; Site:561; Site:562; Site:563; Site:564; Site:565; Site:566; Site:567; Site:568; Site:569; Site:570; Site:571; Site:572; Site:573; Site:574; Site:575; Site:576; Site:577; Site:578; Site:579; Site:580; Site:581; Site:582; Site:583; Site:584; Site:585; Site:586; Site:587; Site:588; Site:589; Site:590; Site:591; Site:592; Site:593; Site:594; Site:595; Site:596; Site:597; Site:598; Site:599; Site:600; Site:601; Site:602; Site:603; Site:604; Site:605; Site:606; Site:607; Site:608; Site:609; Site:610; Site:611; Site:612; Site:613; Site:614; Site:615; Site:616; Site:617; Site:618; Site:619; Site:620; Site:621; Site:622; Site:623; Site:624; Site:625; Site:626; Site:627; Site:628; Site:629; Site:630; Site:631; Site:632; Site:633; Site:634; Site:635; Site:636; Site:637; Site:638; Site:639; Site:640; Site:641; Site:642; Site:643; Site:644; Site:645; Site:646; Site:647; Site:648; Site:649; Site:650; Site:651; Site:652; Site:653; Site:654; Site:655; Site:656; Site:657; Site:658; Site:659; Site:660; Site:661; Site:662; Site:663; Site:664; Site:665; Site:666; Site:667; Site:668; Site:669; Site:670; Site:671; Site:672; Site:673; Site:674; Site:675; Site:676; Site:677; Site:678; Site:679; Site:680; Site:681; Site:682; Site:683; Site:684; Site:685; Site:686; Site:687; Site:688; Site:689; Site:690; Site:691; Site:692; Site:693; Site:694; Site:695; Site:696; Site:697; Site:698; Site:699; Site:700; Site:701; Site:702; Site:703; Site:704; Site:705; Site:706; Site:707; Site:708; Site:709; Site:710; Site:711; Site:712; Site:713; Site:714; Site:715; Site:716; Site:717; Site:718; Site:719; Site:720; Site:721; Site:722; Site:723; Site:724; Site:725; Site:726; Site:727; Site:728; Site:729; Site:730; Site:731; Site:732; Site:733; Site:734; Site:735; Site:736; Site:737; Site:738; Site:739; Site:740; Site:741; Site:742; Site:743; Site:744; Site:745; Site:746; Site:747; Site:748; Site:749; Site:750; Site:751; Site:752; Site:753; Site:754; Site:755; Site:756; Site:757; Site:758; Site:759; Site:760; Site:761; Site:762; Site:763; Site:764; Site:765; Site:766; Site:767; Site:768; Site:769; Site:770; Site:771; Site:772; Site:773; Site:774; Site:775; Site:776; Site:777; Site:778; Site:779; Site:780; Site:781; Site:782; Site:783; Site:784; Site:785; Site:786; Site:787; Site:788; Site:789; Site:790; Site:791; Site:792; Site:793; Site:794; Site:795; Site:796; Site:797; Site:798; Site:799; Site:800; Site:801; Site:802; Site:803; Site:804; Site:805; Site:806; Site:807; Site:808; Site:809; Site:810; Site:811; Site:812; Site:813; Site:814; Site:815; Site:816; Site:817; Site:818; Site:819; Site:820; Site:821; Site:822; Site:823; Site:824; Site:825; Site:826; Site:827; Site:828; Site:829; Site:830; Site:831; Site:832; Site:833; Site:834; Site:835; Site:836; Site:837; Site:838; Site:839; Site:840; Site:841; Site:842; Site:843; Site:844; Site:845; Site:846; Site:847; Site:848; Site:849; Site:850; Site:851; Site:852; Site:853; Site:854; Site:855; Site:856; Site:857; Site:858; Site:859; Site:860; Site:861; Site:862; Site:863; Site:864; Site:865; Site:866; Site:867; Site:868; Site:869; Site:870; Site:871; Site:872; Site:873; Site:874; Site:875; Site:876; Site:877; Site:878; Site:879; Site:880; Site:881; Site:882; Site:883; Site:884; Site:885; Site:886; Site:887; Site:888; Site:889; Site:890; Site:891; Site:892; Site:893; Site:894; Site:895; Site:896; Site:897; Site:898; Site:899; Site:900; Site:901; Site:902; Site:903; Site:904; Site:905; Site:906; Site:907; Site:908; Site:909; Site:910; Site:911; Site:912; Site:913; Site:914; Site:915; Site:916; Site:917; Site:918; Site:919; Site:920; Site:921; Site:922; Site:923; Site:924; Site:925; Site:926; Site:927; Site:928; Site:929; Site:930; Site:931; Site:932; Site:933; Site:934; Site:935; Site:936; Site:937; Site:938; Site:939; Site:940; Site:941; Site:942; Site:943; Site:944; Site:945; Site:946; Site:947; Site:948; Site:949; Site:950; Site:951; Site:952; Site:953; Site:954; Site:955; Site:956; Site:957; Site:958; Site:959; Site:960; Site:961; Site:962; Site:963; Site:964; Site:965; Site:966; Site:967; Site:968; Site:969; Site:970; Site:971; Site:972; Site:973; Site:974; Site:975; Site:976; Site:977; Site:978; Site:979; Site:980; Site:981; Site:982; Site:983; Site:984; Site:985; Site:986; Site:987; Site:988; Site:989; Site:990; Site:991; Site:992; Site:993; Site:994; Site:995; Site:996; Site:997; Site:998; Site:999; Site:1000; Site:1001; Site:1002; Site:1003; Site:1004; Site:1005; Site:1006; Site:1007; Site:1008; Site:1009; Site:1010; Site:1011; Site:1012; Site:1013; Site:1014; Site:1015; Site:1016; Site:1017; Site:1018; Site:1019; Site:1020; Site:1021; Site:1022; Site:1023; Site:1024; Site:1025; Site:1026; Site:1027; Site:1028; Site:1029; Site:1030; Site:1031; Site:1032; Site:1033; Site:1034; Site:1035; Site:1036; Site:1037; Site:1038; Site:1039; Site:1040; Site:1041; Site:1042; Site:1043; Site:1044; Site:1045; Site:1046; Site:1047; Site:1048; Site:1049; Site:1050; Site:1051; Site:1052; Site:1053; Site:1054; Site:1055; Site:1056; Site:1057; Site:1058; Site:1059; Site:1060; Site:1061; Site:1062; Site:1063; Site:1064; Site:1065; Site:1066; Site:1067; Site:1068; Site:1069; Site:1070; Site:1071; Site:1072; Site:1073; Site:1074; Site:1075; Site:1076; Site:1077; Site:1078; Site:1079; Site:1080; Site:1081; Site:1082; Site:1083; Site:1084; Site:1085; Site:1086; Site:1087; Site:1088; Site:1089; Site:1090; Site:1091; Site:1092; Site:1093; Site:1094; Site:1095; Site:1096; Site:1097; Site:1098; Site:1099; Site:1100; Site:1101; Site:1102; Site:1103; Site:1104; Site:1105; Site:1106; Site:1107; Site:1108; Site:1109; Site:1110; Site:1111; Site:1112; Site:1113; Site:1114; Site:1115; Site:1116; Site:1117; Site:1118; Site:1119; Site:1120; Site:1121; Site:1122; Site:1123; Site:1124; Site:1125; Site:1126; Site:1127; Site:1128; Site:1129; Site:1130; Site:1131; Site:1132; Site:1133; Site:1134; Site:1135; Site:1136; Site:1137; Site:1138; Site:1139; Site:1140; Site:1141; Site:1142; Site:1143; Site:1144; Site:1145; Site:1146; Site:1147; Site:1148; Site:1149; Site:1150; Site:1151; Site:1152; Site:1153; Site:1154; Site:1155; Site:1156; Site:1157; Site:1158; Site:1159; Site:1160; Site:1161; Site:1162; Site:1163; Site:1164; Site:1165; Site:1166; Site:1167; Site:1168; Site:1169; Site:1170; Site:1171; Site:1172; Site:1173; Site:1174; Site:1175; Site:1176; Site:1177; Site:1178; Site:1179; Site:1180; Site:1181; Site:1182; Site:1183; Site:1184; Site:1185; Site:1186; Site:1187; Site:1188; Site:1189; Site:1190; Site:1191; Site:1192; Site:1193; Site:1194; Site:1195; Site:1196; Site:1197; Site:1198; Site:1199; Site:1200; Site:1201; Site:1202; Site:1203; Site:1204; Site:1205; Site:1206; Site:1207; Site:1208; Site:1209; Site:1210; Site:1211; Site:1212; Site:1213; Site:1214; Site:1215; Site:1216; Site:1217; Site:1218; Site:1219; Site:1220; Site:1221; Site:1222; Site:1223; Site:1224; Site:1225; Site:1226; Site:1227; Site:1228; Site:1229; Site:1230; Site:1231; Site:1232; Site:1233; Site:1234; Site:1235; Site:1236; Site:1237; Site:1238; Site:1239; Site:1240; Site:1241; Site:1242; Site:1243; Site:1244; Site:1245; Site:1246; Site:1247; Site:1248; Site:1249; Site:1250; Site:1251; Site:1252; Site:1253; Site:1254; Site:1255; Site:1256; Site:1257; Site:1258; Site:1259; Site:1260; Site:1261; Site:1262; Site:1263; Site:1264; Site:1265; Site:1266; Site:1267; Site:1268; Site:1269; Site:1270; Site:1271; Site:1272; Site:1273; Site:1274; Site:1275; Site:1276; Site:1277; Site:1278; Site:1279; Site:1280; Site:1281; Site:1282; Site:1283; Site:1284; Site:1285; Site:1286; Site:1287; Site:1288; Site:1289; Site:1290; Site:1291; Site:1292; Site:1293; Site:1294; Site:1295; Site:1296; Site:1297; Site:1298; Site:1299; Site:1300; Site:1301; Site:1302; Site:1303; Site:1304; Site:1305; Site:1306; Site:1307; Site:1308; Site:1309; Site:1310; Site:1311; Site:1312; Site:1313; Site:1314; Site:1315; Site:1316; Site:1317; Site:1318; Site:1319; Site:1320; Site:1321; Site:1322; Site:1323; Site:1324; Site:1325; Site:1326; Site:1327; Site:1328; Site:1329; Site:1330; Site:1331; Site:1332; Site:1333; Site:1334; Site:1335; Site:1336; Site:1337; Site:1338; Site:1339; Site:1340; Site:1341; Site:1342; Site:1343; Site:1344; Site:1345; Site:1346; Site:1347; Site:1348; Site:1349; Site:1350; Site:1351; Site:1352; Site:1353; Site:1354; Site:1355; Site:1356; Site:1357; Site:1358; Site:1359; Site:1360; Site:1361; Site:1362; Site:1363; Site:1364; Site:1365; Site:1366; Site:1367; Site:1368; Site:1369; Site:1370; Site:1371; Site:1372; Site:1373; Site:1374; Site:1375; Site:1376; Site:1377; Site:1378; Site:1379; Site:1380; Site:1381; Site:1382; Site:1383; Site:1384; Site:1385; Site:1386; Site:1387; Site:1388; Site:1389; Site:1390; Site:1391; Site:1392; Site:1393; Site:1394; Site:1395; Site:1396; Site:1397; Site:1398; Site:1399; Site:1400; Site:1401; Site:1402; Site:1403; Site:1404; Site:1405; Site:1406; Site:1407; Site:1408; Site:1409; Site:1410; Site:1411; Site:1412; Site:1413; Site:1414; Site:1415; Site:1416; Site:1417; Site:1418; Site:1419; Site:1420; Site:1421; Site:1422; Site:1423; Site:1424; Site:1425; Site:1426; Site:1427; Site:1428; Site:1429; Site:1430; Site:1431; Site:1432; Site:1433; Site:1434; Site:1435; Site:1436; Site:1437; Site:1438; Site:1439; Site:1440; Site:1441; Site:1442; Site:1443; Site:1444; Site:1445; Site:1446; Site:1447; Site:1448; Site:1449; Site:1450; Site:1451; Site:1452; Site:1453; Site:1454; Site:1455; Site:1456; Site:1457; Site:1458; Site:1459; Site:1460; Site:1461; Site:1462; Site:1463; Site:1464; Site:1465; Site:1466; Site:1467; Site:1468; Site:1469; Site:1470; Site:1471; Site:1472; Site:1473; Site:1474; Site:1475; Site:1476; Site:1477; Site:1478; Site:1479; Site:1480; Site:1481; Site:1482; Site:1483; Site:1484; Site:1485; Site:1486; Site:1487; Site:1488; Site:1489; Site:1490; Site:1491; Site:1492; Site:1493; Site:1494; Site:1495; Site:1496; Site:1497; Site:1498; Site:1499; Site:1500; Site:1501; Site:1502; Site:1503; Site:1504; Site:1505; Site:1506; Site:1507; Site:1508; Site:1509; Site:1510; Site:1511; Site:1512; Site:1513; Site:1514; Site:1515; Site:1516; Site:1517; Site:1518; Site:1519; Site:1520; Site:1521; Site:1522; Site:1523; Site:1524; Site:1525; Site:1526; Site:1527; Site:1528; Site:1529; Site:1530; Site:1531; Site:1532; Site:1533; Site:1534; Site:1535; Site:1536; Site:1537; Site:1538; Site:1539; Site:1540; Site:1541; Site:1542; Site:1543; Site:1544; Site:1545; Site:1546; Site:1547; Site:1548; Site:1549; Site:1550; Site:1551; Site:1552; Site:1553; Site:1554; Site:1555; Site:1556; Site:1557; Site:1558; Site:1559; Site:1560; Site:1561; Site:1562; Site:1563; Site:1564; Site:1565; Site:1566; Site:1567; Site:1568; Site:1569; Site:1570; Site:1571; Site:1572; Site:1573; Site:1574; Site:1575; Site:1576; Site:1577; Site:1578; Site:1579; Site:1580; Site:1581; Site:1582; Site:1583; Site:1584; Site:1585; Site:1586; Site:1587; Site:1588; Site:1589; Site:1590; Site:1591; Site:1592; Site:1593; Site:1594; Site:1595; Site:1596; Site:1597; Site:1598; Site:1599; Site:1600; Site:1601; Site:1602; Site:1603; Site:1604; Site:1605; Site:1606; Site:1607; Site:1608; Site:1609; Site:1610; Site:1611; Site:1612; Site:1613; Site:1614; Site:1615; Site:1616; Site:1617; Site:1618; Site:1619; Site:1620; Site:1621; Site:1622; Site:1623; Site:1624; Site:1625; Site:1626; Site:1627; Site:1628; Site:1629; Site:1630; Site:1631; Site:1632; Site:1633; Site:1634; Site:1635; Site:1636; Site:1637; Site:1638; Site:1639; Site:1640; Site:1641; Site:1642; Site:1643; Site:1644; Site:1645; Site:1646; Site:1647; Site:1648; Site:1649; Site:1650; Site:1651; Site:1652; Site:1653; Site:1654; Site:1655; Site:1656; Site:1657; Site:1658; Site:1659; Site:1660; Site:1661; Site:1662; Site:1663; Site:1664; Site:1665; Site:1666; Site:1667; Site:1668; Site:1669; Site:1670; Site:1671; Site:1672; Site:1673; Site:1674; Site:1675; Site:1676; Site:1677; Site:1678; Site:1679; Site:1680; Site:1681; Site:1682; Site:1683; Site:1684; Site:1685; Site:1686; Site:1687; Site:1688; Site:1689; Site:1690; Site:1691; Site:1692; Site:1693; Site:1694; Site:1695; Site:1696; Site:1697; Site:1698; Site:1699; Site:1700; Site:1701; Site:1702; Site:1703; Site:1704; Site:1705; Site:1706; Site:1707; Site:1708; Site:1709; Site:1710; Site:1711; Site:1712; Site:1713; Site:1714; Site:1715; Site:1716; Site:1717; Site:1718; Site:1719; Site:1720; Site:1721; Site:1722; Site:1723; Site:1724; Site:1725; Site:1726; Site:1727; Site:1728; Site:1729; Site:1730; Site:1731; Site:1732; Site:1733; Site:1734; Site:1735; Site:1736; Site:1737; Site:1738; Site:1739; Site:1740; Site:1741; Site:1742; Site:1743; Site:1744; Site:1745; Site:1746; Site:1747; Site:1748; Site:1749; Site:1750; Site:1751; Site:1752; Site:1753; Site:1754; Site:1755; Site:1756; Site:1757; Site:1758; Site:1759; Site:1760; Site:1761; Site:1762; Site:1763; Site:1764; Site:1765; Site:1766; Site:1767; Site:1768; Site:1769; Site:1770; Site:1771; Site:1772; Site:1773; Site:1774; Site:1775; Site:1776; Site:1777; Site:1778; Site

THIS PAGE BLANK (USPTO)

XX	Sanders MC.
PI	
XX	WPI; 2002-303940/34.
DR	
XX	Detection of bacterial contamination in foods or food-related work
PT	areas comprises identifying a protein specific to the microorganism
PT	using fluorescence or colorimetric methods
XX	
PS	Disclosure; Page 7; 25pp; English.
XX	
CC	The invention describes a specific and sensitive method for detection of
CC	a pathogenic microorganism in potentially contaminated food products
CC	at the retail level, by detecting a bacteria-specific protein e.g. a
CC	Listeria monocytogenes metalloprotease. The process is useful for
CC	detecting Listeria monocytogenes, a common cause of gastroenteritis, or
CC	other food contaminants in food products or food-related work areas.
CC	This sequence represents the peptide P2 that functions as a control
CC	substrate in fluorescence resonance energy transfer (FRET) analysis of
CC	bacterial protease activity in food products.
XX	
SO	Sequence 11 AA;
	Query Match 100.0%; Score 56; DB 23; Length 11;
	Best Local Similarity 100.0%; Pred. No. 0.00046;
	Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY	1 ADTVEPTGAKE 11
Db	1 ADTVEPTGAKE 11
RESULT 2	
AAW97108	
ID	AAW97108 standard; Protein; 479 AA.
XX	
AC	AAW97108;
XX	
DT	29-APR-1999 (first entry)
XX	
DE	Thermostable polypeptide factor binding to DNA polymerase.
XX	
KW	Thermostable polypeptide factor; DNA synthesis activity;
KW	DNA polymerase; in vitro DNA synthesis.
XX	
OS	Pyrococcus furiosus.
PN	WO9900506-A1.
XX	
PD	07-JAN-1999.
XX	
PF	24-JUN-1998; 98WO-JP02845.
XX	
PR	21-NOV-1997; 97JP-0320692.
PR	26-JUN-1997; 97JP-0187496.
XX	
PA	(TAKI) TAKARA SHUZO CO LTD.
XX	
PI	Asada K, Fujita T, Kato I, Miyake K, Mukai H, Sato Y;
PI	Uemori T;
XX	
DR	WPI; 1999-095751/08.
XX	
XX	N-PSDB; AAX15304.
PT	
PT	Thermostable polypeptide factors promoting the activity of DNA
PT	polymerase - for improvement of DNA synthesis and amplification in
PT	vitro.
XX	
PS	Claim 4; Page 134-137; 177pp; Japanese.
XX	
CC	The present sequence represents a thermostable polypeptide factor
CC	which binds to, and promotes the DNA synthesis activity of DNA
CC	polymerase. The polymerase related factors can be used to provide

CC	more efficient in vitro DNA synthesis and amplification systems
CC	(e.g. for polymerase chain reaction) by using the factors in
CC	conjunction with a DNA polymerase.
XX	
SQ	Sequence 479 AA;
OY	Query Match 82.1%; Score 46; DB 20; Length 479; Best Local Similarity 72.7%; Pred. No. 2.1; Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0
DB	1 ADVEPTGAKK 11 : 110 ADNIEPSAKE 120
RESULT 3	
AAB62022	AAB62022 standard; Protein; 479 AA.
AC	AAB62022;
XX	
DT	14-MAY-2001 (first entry)
DE	P. furiosus recombinant P55 amino acid sequence.
XX	
KW	MCM; minichromosome maintenance protein; archaeal polypeptide; PCNA; Ref-P38; Rfc-P55; RFA; CDC6; FEN-1; dTTPase; Ligase; helicase dna2; PCR; helicase 4; nucleic acid amplification; polymerase chain reaction.
XX	
OS	Pyrococcus furiosus.
XX	
FH	Key Location/Qualifiers
FT	Misc-difference 354 /label= unknown
FT	/note="encoded by "
XX	
PN	MO200109347-A2.
XX	
PD	08-FEB-2001.
XX	
PF	28-JUL-2000; 2000MO-U05032.
XX	
PR	30-JUL-1999; 99US-0146580.
XX	
PA	(STRA-) STRATAGENE.
XX	
PI	Hogrefe HH, Cline JM, Hansen CJ, Borns MC;
DR	WPL; 2001-182959/18.
XX	
N-PSDB:	AAE57027.
XX	
PT	Composition for improving nucleic acid polymerase reactions, useful
PT	e.g. in synthesis or amplification, contains at least one archaeal
XX	accessory protein -
PS	Disclosure; Fig 10; 147pp; English.
XX	
XX	The invention provides a composition (A) for enhancing nucleic acid
CC	polymerase reactions that comprises an archaeal MCM (minichromosome
CC	maintenance protein) and at least one of the archaeal polypeptides (PCNA,
CC	RFC-P38 or -P55, RFA, CDC6, FEN-1, dTTPase, ligase, helicase dna2, or
CC	helicases 2-8); (A) And similar compositions containing different
CC	combinations of accessory proteins, are used to improve performance of
CC	synthesis, amplification, mutagenizing, labeling and detecting reactions,
CC	e.g. for gene characterization, cloning, detection of allelic variants,
CC	diagnosis and screening for disease, particularly where done by
CC	polymerase chain reaction (PCR). Some of the proteins also stabilize
CC	duplicates during polymerase reactions or improve exonuclease reactions,
CC	for example RFA also improves specificity of nucleic acid/protein
CC	interaction and PCNA improves polymerase-mediated repair processes and
CC	hybridization reactions. Nucleic acids encoding the archaeal polypeptides
CC	are used for recombinant production of proteins, and fragments of the
CC	nucleic acids as probes and primers for screening related sequences. The

CC accessory proteins increase accuracy and efficiency of polymerase
CC reactions, allow use of lower denaturation and extension temperatures.
CC (possibly isothermal processing), and improve synthesis of long targets.
CC The present sequence represents the amino acid sequence of *P. furiosus*
CC recombinant p55.

XX Sequence 479 AA;

Query Match 82.1%; Score 46; DB 22; Length 479;

Best Local Similarity 72.7%; Pred. No. 2.1;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 ADTVEPTGAKE 11
11:11:1111

Db 110 ADNIEPSGAKE 120

RESULT 4

AB62021 AAB62021 standard; Protein; 1615 AA.

XX AAB62021;

DT 14-MAY-2001 (first entry)

DE *P. furiosus* genomic RFC amino acid sequence.

XX MCM: minichromosome maintenance protein; archaeal polypeptide; PCNA;

KW RFC-P38; RFC-P55; RFA; CDC6; FEN-1; dUTPase; Ligase; helicase dna2;

KM helicase; nucleic acid amplification; polymerase chain reaction; PCR.

OS *Pyrococcus furiosus*.

FT Key Location/Qualifiers
FT Misc-difference 1..1615
FT /note="Xaa are residues encoded by degenerate
FT codons and/or stop codons"

PN WO200109347-A2.

PD 08-FEB-2001.

PF 28-JUL-2000; 2000WO-US20532.

PR 30-JUL-1999; 95US-0146580.

PA (STRA-) STRATAGENE.

PI Hogrefe HH, Cline JM, Hansen CJ, Borns MC;

DR MPI: 2001-182959/18.

DR N-PSDB; AAF57027.

PT Composition for improving nucleic acid polymerase reactions, useful
PT e.g. in synthesis or amplification, contains at least one archaeal
PT accessory protein -

PS Disclosure; Fig 9; 147pp; English.

XX The invention provides a composition (A) for enhancing nucleic acid
CC polymerase reactions that comprises an archaeal MCM (minichromosome
CC maintenance protein) and at least one of the archaeal polypeptides (PCNA,
CC RFC-P38 or -P55, RFA, CDC6, FEN-1, dUTPase, ligase, helicase dna2, or
CC helicases 2-8). (A) And similar compositions containing different
CC combinations of accessory proteins, are used to improve performance of
CC syntheses, amplification, mutagenizing, labeling and detecting reactions,
CC e.g. for gene characterization, cloning, detection of allelic variants,
CC diagnosis and screening for disease, particularly where done by
CC polymerase chain reaction (PCR). Some of the proteins also stabilize
CC duplexes during polymerase reactions or improve exonuclease reactions,
CC for example RFA also improves specificity of nucleic acid/protein
CC interaction and PCNA improves polymerase-mediated repair processes and
CC hybridization reactions. Nucleic acids encoding the archaeal polypeptides

CC are used for recombinant production of proteins, and fragments of the
CC nucleic acid as probes and primers for screening related sequences. The
CC accessory proteins increase accuracy and efficiency of polymerase
CC reactions, allow use of lower denaturation and extension temperatures.
CC (possibly isothermal processing), and improve synthesis of long targets.
CC The present sequence represents the sequence of *P. furiosus* genomic
CC RFC, including the p38 and p55 subunits.

XX Sequence 1615 AA;

Query Match 82.1%; Score 46; DB 22; Length 1615;

Best Local Similarity 72.7%; Pred. No. 7.8;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 ADTVEPTGAKE 11
11:11:1111

Db 1030 ADNIEPSGAKE 1040

RESULT 5

AB647676 ABB47676 standard; Protein; 289 AA.

XX ABB47676;

DT 05-FEB-2002 (first entry)

DE *Listeria monocytogenes* protein #380.

XX Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;

KW vitamin B12; bacterial infection; disease.

XX *Listeria monocytogenes*.

OS *Listeria monocytogenes*.

PN WO200177335-A2.

PD 18-OCT-2001.

PF 11-APR-2001; 2001WO-FR01118.

PR 11-APR-2000; 2000FR-0004629.

PA (INSP) INSP PASTEUR.

PI Buchrieser C, Frangoul L, Couve E, Rusnick C, Fsihi H, Dehoux P;

PI Dussurget O, Chetoui F, Nedjari H, Glaser P, Kunst F, Cossart P;

PI Daniels J, Goebel W, Kreft J, Kunz M, Ng E, Vazquez-Boland JA;

PI Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A;

PI Chakraborty T, Domann E, Hain T, Berche P, Chardit A, Durant L;

PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N;

PI Madueno E, De Pablos B, Wehland J, Kaerst U, Ertlan K, Hauf J;

DR MPI: 2002-010914/01.

PT Genomic sequence for *Listeria monocytogenes*, useful e.g. for treatment
PT and prevention of *Listeria* and related bacterial infections, and
PT related polypeptides -

PS Claim 6; SEQ ID No 381; 192pp; French.

XX The present invention relates to the genome sequence of *Listeria*
CC monocytogenes EGD-e (see ABA03041). The genome sequence and fragments of
CC it are useful for selecting probes and primers for detecting genes in *L.*
CC *monocytogenes* and related organisms, and for studying genetic
CC polymorphisms and other genomes. The present sequence is a protein
CC encoded by the genome sequence of the present invention. Proteins
CC expressed from the genome sequence are useful for raising specific
CC antibodies, identification of *L. monocytogenes* and related organisms,
CC for biosynthesis and biodegradation, especially biosynthesis of Vitamin
CC B12. The genome sequence and proteins encoded by it are also useful for
CC selecting compounds that regulate gene expression and cell replication
CC and modulate *L. monocytogenes*-related diseases. In addition, the genome

QY 1 ADTVEPTGAKE 11

DE P. horikoshii PHBN013 protein fragment
YY

KW Thermostable; template-dependent elongation; staple protein;
KM elongation protein; amplification; reverse transcription.
XX
OS Pyrococcus horikoshii.
XX
PN WO200008164-A2.
XX
PD 17-FEB-2000.
XX
PF 06-AUG-1999; 99WO-DE02480.
XX
PR 06-AUG-1998; 98DE-1035653.
PR 07-SEP-1998; 98DE-1040771.
PR 18-JUN-1999; 99EP-0111795.
XX
PA (LION-) LION BIOSCIENCE AG.
XX
PI Kilger C, Kober I, Voss H, Moeckel G;
DR WPI; 2000-195576/17.
XX
PT Thermostable in vitro polymerase complex for template-dependent
PT elongation of nucleic acids -
PS Disclosure; Page 136-138; 233pp; German.
XX
XX This invention describes a novel thermostable in vitro complex (I) for
CC template-dependent elongation of nucleic acids (NA) comprising a
CC thermostable 'staple' protein and an elongation protein. The thermostable
CC in vitro complex is useful for template-dependent elongation of NA, e.g.
CC for amplification or reverse transcription. This is useful for sequencing
CC NA by the polymerase chain reaction or reverse transcriptase-PCR. The
CC complex can be used to mark NA. This sequence represents the
CC Pyrococcus horikoshii PHB013 protein which is used to illustrate the
CC method of the invention.
XX
SQ Sequence 468 AA;

Query Match 71.4%; Score 40; DB 21; Length 468;
Best Local Similarity 63.6%; Pred. No. 28;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 ADTVEPTGAKE 11
II :||:| | |
Db 110 ADNIEPSGAPE 120

RESULT 9
AAB96068
ID AAB96068 standard; Protein; 479 AA.
XX
AC AAB96068;
XX
DT 29-OCT-2001 (first entry)
XX
DE Putative P. abyssi replication factor C #2.
XX
KM Hyperthermophilic archaeon; hyperthermophilic protein.
XX
OS Pyrococcus abyssi.
XX
PN FR2792651-A1.
XX
PD 27-OCT-2000.
XX
PE 21-APR-1999; 99FR-0005034.
XX
PR 21-APR-1999; 99FR-0005034.
XX
PA (CNRS) CNRS CENT NAT RECH SCI.
PA (IFRE-) IFREMER INST FR RECH EXPL MER.
XX
PI Forterre P, Thierry JC, Prieur D, Dietrich J, Lecompte O;

PI Querellou J, Weissenbach J, Saurin W, Heilig R;
XX
DR WPI; 2001-126236/14.
XX
PT New nucleotide sequences isolated from Pyrococcus abyssi encode
PT proteins useful in industry -
XX
PS Claim 7; Pages 687-689; 1657pp; French.
XX
CC The present invention relates to the genomic sequence of Pyrococcus
CC abyssi (see AAF86431 and AAH41223-7) and P. abyssi proteins. P. abyssi is
CC a hyperthermophilic archaeon, which is isolated from deep-sea
CC hydrothermal vents. The present sequence is one such P. abyssi protein.
CC The proteins of the present invention have various potential industrial
CC uses, since the proteins are stable at very high temperatures, some up to
CC 110 degrees centigrade.
CC Note: This patent is in the same patent family as WO200065062, which
CC contains additional sequences as shown in AAB99132-AAB99143,
CC AAH75903-AAH75920 and AAG66436.
XX
SQ Sequence 479 AA;

Query Match 71.4%; Score 40; DB 22; Length 479;
Best Local Similarity 63.6%; Pred. No. 29;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 ADTVEPTGAKE 11
II :||:| | |
Db 110 ADNIEPSGAPE 120

RESULT 10
AAG90944
ID AAG90944 standard; Protein; 395 AA.
XX
AC AAG90944;
XX
DT 26-SEP-2001 (first entry)
XX
DE C glutamicum protein fragment SEQ ID NO: 4698.
XX
KM Corynebacterium; amino acid synthesis; vitamin; saccharide;
KM organic acid synthesis.
XX
OS Corynebacterium glutamicum.
XX
PN EP1108790-A2.
XX
PD 20-JUN-2001.
XX
PF 18-DEC-2000; 2000EP-0127688.
XX
PR 16-DEC-1999; 99JP-0377484.
PR 07-APR-2000; 2000JP-0159162.
PR 03-AUG-2000; 2000JP-0280988.
XX
PA (KYOWA) KYOWA HAKKO KOGYO KK.
XX
PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
XX
DR WPI; 2001-376931/40.
XX
DR N-PSDB; AAH66163.
XX
PT Novel polynucleotides derived from Corynebacterium bacteria, for identifying
PT mutation point of a gene, measuring expression of a gene, analysing
PT expression profile or pattern of a gene and identifying homologous gene
XX
PS Claim 17; SEQ ID NO: 4698; 246pp + Sequence Listing; English.
XX
CC The present invention provides a number of nucleotide and protein
CC sequences from the Corynebacterium glutamicum. These

CC are useful for identifying the mutation point of a gene derived from a
CC mutant of *Corynebacterium*, measuring expression amount and
CC analysing the expression profile or expression pattern of a gene derived
CC from *Corynebacterium*, and identifying a homolog of a gene derived
CC from *Corynebacterium*. *Corynebacterium* bacteria are useful for producing
CC amino acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a protein described
CC in the exemplification of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC European Patent Office.

XX Sequence 395 AA;

Query Match 67.9%; Score 38; DB 22; Length 395;

Best Local Similarity 80.0%; Pred. No. 57;

Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 ADTVEPTGAK 10
II :||||:|

DB 288 AHTVPTGAK 297

RESULT 11

AAU51333

ID AAU51333 standard; Protein; 830 AA.

XX AC AAU51333;

DT 27-FEB-2002 (first entry)

DE Propionibacterium *acnes* immunogenic protein #12229.

XX SAPHO syndrome; synovitis; acne; pustulosis; hyperostosis; osteomyelitis;
KM uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KM inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
XX dermatological; osteopathic; neuroprotectant.

XX Propionibacterium *acnes*.

XX PN WO200181581-A2.

XX PD 01-NOV-2001.

XX PF 20-APR-2001; 2001WO-US12865.

XX PR 21-APR-2000; 2000US-199047P.

XX PR 02-JUN-2000; 2000US-208841P.

XX PR 07-JUL-2000; 2000US-216747P.

XX PA (CORI-) CORIXA CORP.

XX PI Skelky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;

XX PI L'maisonneuve J, Zhang Y, Jen S, Carter D;

XX DR WPI: 2001-616774/71.

XX DR N-PSDB: AAS59551.

XX PT Propionibacterium *acnes* polypeptides and nucleic acids useful for

XX PT vaccinating against and diagnosing infections, especially useful for

XX PT treating acne vulgaris -

XX PS Example 1; SEQ ID No 12528; 1069pp; English.

XX Sequences AAU59105-AAU68017 represent *Propionibacterium acnes* immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC *P. acnes*. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.
CC *P. acnes* is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of *P. acnes* in a patient comprises contacting a

CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for *P. acnes* proteins. These antibodies can be used to
CC downregulate expression and activity of *P. acnes* polypeptides and
CC therefore treat *P. acnes* infections. The antibodies may also be used as
CC diagnostic agents for determining *P. acnes* presence, for example, by
CC enzyme linked immunosorbent assay (ELISA).
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pt_sequences.

XX Sequence 830 AA;

Query Match 67.9%; Score 38; DB 22; Length 830;

Best Local Similarity 63.6%; Pred. No. 1.3e+02;

Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 ADTVEPTGAK 11
II :||||:|

DB 586 ADWDPPTGARE 596

RESULT 12

AAG98333

ID AAG98333 standard; Protein; 236 AA.

XX AC AAG98333;

DT 21-SEP-2001 (first entry)

DE *Escherichia coli* protein sequence SEQ ID NO:381.

XX *Escherichia coli*; identification; proliferation; microorganism;
KM antimicrobial; antibacterial; antibiotic; gene therapy; diagnosis;
KM bacterial growth inhibition.

XX OS *Escherichia coli*.

XX PN WO200148209-A2.

XX PD 05-JUL-2001.

XX PF 19-DEC-2000; 2000WO-US34419.

XX PR 23-DEC-1999; 99US-0173005.

XX PA (ELIT-) ELITRA PHARM INC.

XX PI Forsyth RA, Ohlsen KL, Zyskind JW;

XX PI WPI: 2001-457376/49.

XX DR N-PSDB: AAH81389.

XX PT Novel nucleic acids encoding proteins required for *Escherichia coli*

XX PT proliferation, useful for screening for antimicrobial agents -

XX PS Claim 19; Page 503-504; 596pp; English.

XX The present invention describes a purified or isolated nucleic acid
CC sequence (I) consisting essentially of one of the 93 nucleotide sequences
CC given in AAH81202 to AAH81294, where expression of the nucleic acid in a
CC microorganism is capable of inhibiting proliferation of a microorganism.
CC (I) have antibacterial and antibiotic activities, and can be used in
CC gene therapy. Expression of (I) in a microorganism inhibits proliferation
CC of the microorganism, and the manufactured antibiotic is useful for
CC reducing the activity or level of a gene product required for
CC proliferation of a microorganism in a subject, specifically humans. The
CC nucleic acids that inhibit bacterial growth or proliferation can be used
CC as antisense therapeutics for killing bacteria. In addition to
CC therapeutic applications, the nucleic acid sequences complementary to
CC sequences required for proliferation can be used as diagnostic tools.
CC For example, nucleic acid probes complementary to proliferation-required

CC sequences that are specific for particular species of microorganisms can
CC be used as probes to identify particular microorganism species in
CC clinical specimens. AAH81295 to AAH81487 encode the *Escherichia coli*
CC proteins given in AAC98239 to AAC98431, and AAH81488 to AAH81491
CC represent oligonucleotides, which are used in the exemplification of the
CC present invention.

XX SQ Sequence 236 AA;

Query Match 66.1%; Score 37; DB 22; Length 236;

Best Local Similarity 80.0%; Pred. No. 51;

Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ADTVEPTGAK 10
| | | | |

DB 69 ADLVEATGAK 78

RESULT 13

AAH95711
ID AAB95711 standard; Protein; 374 AA.

XX AC AAB95711;

XX DT 26-JUN-2001 (first entry)

XX DE Human protein sequence SEQ ID NO:18558.

XX KM Human; primer; detection; diagnosis; antisense therapy; gene therapy.

XX OS Homo sapiens.

XX PN EP1074617-A2.

XX PD 07-FEB-2001.

XX PF 28-JUL-2000; 2000EP-0116126.

XX PR 29-JUL-1999; 99JP-0248036.

XX PR 27-AUG-1999; 99JP-0300253.

XX PR 11-JAN-2000; 2000JP-0118776.

XX PR 02-MAY-2000; 2000JP-0183767.

XX PR 09-JUN-2000; 2000JP-0241899.

XX PA (HELI-) HELIX RES INST.

XX PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

XX PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX DR WPI; 2001-318749/34.

XX PS Claim 8; SEQ ID 18558; 2537pp + CD ROM; English.

XX The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC -oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the

CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.

XX SQ Sequence 374 AA;

Query Match 66.1%; Score 37; DB 22; Length 374;

Best Local Similarity 100.0%; Pred. No. 83;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 EPTGAK 11
| | | | |

DB 6 EPTGAK 12

RESULT 14

AAU36361
ID AAU36361 standard; Protein; 459 AA.

XX AC AAU36361;

XX DT 14-FEB-2002 (first entry)

XX DE Pseudomonas aeruginosa cellular proliferation protein #351.

XX KM Antisense; prokaryotic cellular proliferation protein;

XX KW antibiotic; antibacterial; drug design.

XX OS Pseudomonas aeruginosa.

XX PN MO200170955-A2.

XX PD 27-SEP-2001.

XX PF 21-MAR-2001; 2001WO-US09180.

XX PR 21-MAR-2000; 2000US-191078P.

XX PR 23-MAY-2000; 2000US-206848P.

XX PR 26-MAY-2000; 2000US-207727P.

XX PR 23-OCT-2000; 2000US-242578P.

XX PR 27-NOV-2000; 2000US-253625P.

XX PR 22-DEC-2000; 2000US-257931P.

XX PR 16-FEB-2001; 2001US-269308P.

XX PA (ELITR-) ELITRA PHARM INC.

XX PI Haselbeck R, Ohlsen KL, Zyskind JM, Wall D, Trawick JD, Carr GJ;

XX PI Yamamoto RT, Xu HH;

XX DR WPI; 2001-611495/70.

XX DR N-PSDB; AAS54220.

XX PT New polynucleotides for the identification and development of

XX PT antibiotics, comprise sequences of antisense nucleic acids -

XX PS Example 3; Seq ID No 11954; 511pp; English.

XX The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella*
CC *pneumoniae*, *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen

CC for homologous nucleic acids which are required for cell proliferation in
 CC a wide variety of organisms. The present sequence represents an
 CC essential prokaryotic cellular proliferation protein.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 459 AA;
 Query Match 66.1%; Score 37; DB 22; Length 459;
 Best Local Similarity 87.5%; Pred. No. 1e+02;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 OY 1 ADTVEPTG 8
 |||||
 Db 105 ADTVEPTAG 112
 RESULT 15
 ABG09857
 ID ABG09857 standard; Protein; 60 AA.
 XX
 AC ABG09857;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #9848.
 XX
 KM Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KM food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 PN MO200175067-A2.
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US08631.
 XX
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Drmanac RT, Liu C, Tang YT;
 XX
 DR WPI; 2001-639362/73.
 DR N-PSDB; AAS74044.
 XX
 XX
 PS Claim 20; SEQ ID NO 40216; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and

CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 60 AA;
 Query Match 64.3%; Score 36; DB 22; Length 60;
 Best Local Similarity 77.8%; Pred. No. 18;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 OY 1 ADTVEPTGA 9
 |||||
 Db 40 ADTLEPWGA 48
 Search completed: November 18, 2002, 15:55:28
 Job time : 33.9355 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 18, 2002, 15:53:18 ; Search time 12.7742 Seconds
(without alignments)
25.336 Million cell updates/sec

Title: US-09-848-781-3
Perfect score: 56
Sequence: 1 ADTVEPTGAKK 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCRTUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	46	82.1	479	4	US-09-446-504-64
2	46	82.1	479	4	US-09-712-266-64
3	34	60.7	604	2	US-08-468-576B-12
4	34	60.7	604	2	US-08-468-579B-12
5	34	60.7	604	3	US-08-468-577B-12
6	33	58.9	8	4	US-08-160-604-105
7	33	58.9	12	4	US-08-160-604-103
8	33	58.9	25	6	5185431-9
9	33	58.9	459	2	US-08-870-518-4
10	33	58.9	1618	1	US-07-853-913-4
11	33	58.9	1805	3	US-07-853-913-2
12	32	57.1	8	4	US-08-160-604-104
13	32	57.1	256	4	US-09-071-035-300
14	32	57.1	284	4	US-09-071-035-298
15	32	57.1	309	1	US-08-236-918A-2
16	32	57.1	309	4	US-09-150-864A-2
17	32	57.1	422	4	US-09-151-102-2
18	32	57.1	422	4	US-08-929-846-2
19	32	57.1	423	4	US-08-702-665A-5
20	31	55.4	227	6	5428012-3
21	31	55.4	227	6	5428012-4
22	31	55.4	227	6	5451506-3
23	31	55.4	228	6	5428012-2
24	31	55.4	228	6	5451506-2
25	31	55.4	252	1	US-08-792-019B-13
26	31	55.4	252	1	US-08-988-819-13
27	31	55.4	252	4	US-09-016-534-13

28	31	55.4	252	4	US-08-097-869-3	Sequence 3, Appl
29	31	55.4	353	1	US-08-176-620A-14	Sequence 14, Appl
30	31	55.4	353	2	US-08-461-985-14	Sequence 14, Appl
31	31	55.4	353	4	US-08-932-787B-19	Sequence 19, Appl
32	31	55.4	353	4	US-08-932-012C-19	Sequence 19, Appl
33	31	55.4	353	4	US-08-888-818C-19	Sequence 19, Appl
34	31	55.4	424	4	US-09-134-001C-3876	Sequence 3876, Ap
35	31	55.4	432	4	US-08-702-665A-3	Sequence 3, Appl
36	31	55.4	441	4	US-09-151-102-4	Sequence 4, Appl
37	31	55.4	441	4	US-08-929-846-4	Sequence 4, Appl
38	31	55.4	548	2	US-09-032-315-4	Sequence 4, Appl
39	31	55.4	548	2	US-08-993-318A-4	Sequence 4, Appl
40	31	55.4	548	4	US-09-399-886-4	Sequence 4, Appl
41	31	55.4	548	4	US-09-396-260-4	Sequence 4, Appl
42	31	55.4	548	4	US-09-576-281-4	Sequence 4, Appl
43	31	55.4	1401	4	US-09-127-670-6	Sequence 6, Appl
44	31	55.4	3443	2	US-08-416-603-2	Sequence 2, Appl
45	30	53.6	48	3	US-08-804-439A-105	Sequence 105, App

ALIGNMENTS

```
RESULT 1
US-09-446-504-64
; Sequence 64, Application US/09446504
; Patent No. 6218150
; GENERAL INFORMATION:
; APPLICANT: UEMORI, Takashi
; APPLICANT: SATO, Yoshimi
; APPLICANT: FUJITA, Tomoko
; APPLICANT: MIYAKE, Kazue
; APPLICANT: MURAI, Hiroyuki
; APPLICANT: ASADA, Kiyozo
; APPLICANT: KATO, Ikunoshin
; TITLE OF INVENTION: DNA POLYMERASE-RELATED FACTORS
; FILE REFERENCE: 1422-408PCT
; CURRENT APPLICATION NUMBER: US/09/446,504
; CURRENT FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: PCT/JP98/02845
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: JP 9-187496
; PRIOR FILING DATE: 1997-06-26
; PRIOR APPLICATION NUMBER: JP 9-320692
; PRIOR FILING DATE: 1997-11-27
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 64
; LENGTH: 479
; TYPE: PRT
; ORGANISM: Pyrococcus furiosus
US-09-446-504-64

Query Match      82.1%; Score 46; DB 4; Length 479;
Best Local Similarity 72.7%; Pred. No. 0.38;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      1 ADTVEPTGAKK 11
Db      110 ADTVEPTGAKK 120

RESULT 2
US-09-712-266-64
; Sequence 64, Application US/09712266
; Patent No. 6333158
; GENERAL INFORMATION:
; APPLICANT: UEMORI, Takashi
; APPLICANT: SATO, Yoshimi
; APPLICANT: FUJITA, Tomoko
; APPLICANT: MIYAKE, Kazue
; APPLICANT: MURAI, Hiroyuki
; APPLICANT: ASADA, Kiyozo
```

APPLICANT: KATO, Ikuoshin
TITLE OF INVENTION: DNA POLYMERASE-RELATED FACTORS
FILE REFERENCE: 1422-408PCT
CURRENT APPLICATION NUMBER: US/09/712,266
CURRENT FILING DATE: 2000-11-15
PRIOR APPLICATION NUMBER: US 09/446,504
PRIOR FILING DATE: 1999-12-23
PRIOR APPLICATION NUMBER: PCT/JP98/02845
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: JP 9-187496
PRIOR FILING DATE: 1997-06-26
PRIOR APPLICATION NUMBER: JP 9-320692
PRIOR FILING DATE: 1997-11-27
NUMBER OF SEQ ID NOS: 92
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 64
LENGTH: 479
TYPE: PRT
ORGANISM: Pyrococcus furiosus
US-09-712-266-64

Query Match 82.1%; Score 46; DB 4; Length 479;
Best Local Similarity 72.7%; Pred. No. 0.38;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 ADVEPTGAK 11
Db 110 ADNIEPSGAK 120

RESULT 3
US-08-468-576B-12
Sequence 12, Application US/08468576B
Patent No. 5955345

GENERAL INFORMATION:

APPLICANT: Rabin, Daniel

TITLE OF INVENTION: PANCREATIC ISLET CELL ANTIGENS

TITLE OF INVENTION: OBTAINED BY MOLECULAR CLONING

NUMBER OF SEQUENCES: 19

CORRESPONDENCE ADDRESS:

ADDRESSEE: Sprung Kramer Schaefer & Briscoe

STREET: 660 White Plains Road

CITY: Tarrytown

STATE: New York

COUNTRY: USA

ZIP: 10591-5144

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4 Mb storage

COMPUTER: Apple Macintosh

OPERATING SYSTEM: System 7.5

SOFTWARE: WordPerfect

APPLICATION NUMBER: US/08/468,576B

FILING DATE: 06-JUN-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/239,276

FILING DATE: 05-MAY-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/872,646

FILING DATE: 08-JUN-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/715,181

FILING DATE: 14-JUN-1991

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/441,703

FILING DATE: 04-DEC-1989

APPLICATION NUMBER: US 07/312,543

FILING DATE: 17-FEB-1989

ATTORNEY/AGENT INFORMATION:

NAME: Kuitl G. Briscoe

REGISTRATION NUMBER: 33,141

REFERENCE/DOCKET NUMBER: MDI 251.7-KGB
TELECOMMUNICATION INFORMATION:
TELEPHONE: (914) 332-1700
TELEFAX: (914) 332-1844
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 604 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-468-576B-12

Query Match 60.7%; Score 34; DB 2; Length 604;
Best Local Similarity 70.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ADVEPTGAK 10
Db 395 AGTVSPGIR 404

RESULT 4
US-08-468-579B-12
Sequence 12, Application US/08468579B
Patent No. 5981700

GENERAL INFORMATION:

APPLICANT: Rabin, Daniel

TITLE OF INVENTION: PANCREATIC ISLET CELL ANTIGENS

TITLE OF INVENTION: OBTAINED BY MOLECULAR CLONING

NUMBER OF SEQUENCES: 19

CORRESPONDENCE ADDRESS:

ADDRESSEE: Sprung Kramer Schaefer & Briscoe

STREET: 660 White Plains Road

CITY: Tarrytown

STATE: New York

COUNTRY: USA

ZIP: 10591-5144

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4 Mb storage

COMPUTER: Apple Macintosh

OPERATING SYSTEM: System 7.5

SOFTWARE: WordPerfect

APPLICATION NUMBER: US/08/468,579B

FILING DATE: 06-JUN-1995

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/239,276

FILING DATE: 05-MAY-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/872,646

FILING DATE: 08-JUN-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/715,181

FILING DATE: 14-JUN-1991

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/441,703

FILING DATE: 04-DEC-1989

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/312,543

FILING DATE: 17-FEB-1989

ATTORNEY/AGENT INFORMATION:

NAME: Kuitl G. Briscoe

REGISTRATION NUMBER: 33,141

REFERENCE/DOCKET NUMBER: MDI 251.5-KGB

TELECOMMUNICATION INFORMATION:

TELEPHONE: (914) 332-1700

TELEFAX: (914) 332-1844

INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:

LENGTH: 604 amino acids

TYPE: amino acid

TOPOLOGY: linear

US-08-468-579B-12

Query Match 60.7%; Score 34; DB 2; Length 604;
Best Local Similarity 70.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ADTVEPTGAK 10
DB 395 AGTVSPTGIK 404

RESULT 5

US-08-468-577B-12
; Sequence 12, Application US/08468577B
; Patent No. 6001804
; GENERAL INFORMATION:
; APPLICANT: Rabin, Daniel
; TITLE OF INVENTION: PANCREATIC ISLET CELL ANTIGENS
; TITLE OF INVENTION: OBTAINED BY MOLECULAR CLONING
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sprung Kramer Schaefer & Briscoe
; STREET: 660 White Plains Road
; CITY: Tarrytown
; STATE: New York
; COUNTRY: USA
; ZIP: 10591-5144
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4 Mb storage
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: System 7.5
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,577B
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/239,276
; FILING DATE: 05-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/872,646
; FILING DATE: 08-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/715,181
; FILING DATE: 14-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/441,703
; FILING DATE: 04-DEC-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/312,543
; FILING DATE: 17-FEB-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Kuit G. Briscoe
; REGISTRATION NUMBER: 33,141
; REFERENCE/DOCKET NUMBER: MDT 251.8-KGB
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (914) 332-1700
; TELEFAX: (914) 332-1844
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 604 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-468-577B-12

Query Match 60.7%; Score 34; DB 3; Length 604;
Best Local Similarity 70.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ADTVEPTGAK 10
DB 395 AGTVSPTGIK 404

RESULT 6
US-08-160-604-105
; Sequence 105, Application US/08160604
; Patent No. 6232522
; GENERAL INFORMATION:
; APPLICANT: Harley, John
; APPLICANT: James, Judith A.
; APPLICANT: Scofield, R. H.
; TITLE OF INVENTION: PEPTIDE INDUCTION OF AUTOIMMUNITY AND CLINICAL SYMPTOMATOLOG
; NUMBER OF SEQUENCES: 127
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 1100 Peachtree Street, Suite 2800
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-4530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/160,604
; FILING DATE: 30-NOV-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/867,819
; FILING DATE: 13-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/648,205
; FILING DATE: 31-JAN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/472,947
; FILING DATE: 31-JAN-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: OMRF114CIP(3)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404)-815-6508
; TELEFAX: (404)-815-6555
; INFORMATION FOR SEQ ID NO: 105:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; STRANDEDNESS: Single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
US-08-160-604-105

Query Match 58.9%; Score 33; DB 4; Length 8;
Best Local Similarity 75.0%; Pred. No. 1.9e+05;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 TVEPTGAK 10
DB 1 TVEPTGKR 8

RESULT 7
US-08-160-604-103
; Sequence 103, Application US/08160604
; Patent No. 6232522
; GENERAL INFORMATION:
; APPLICANT: Harley, John
; APPLICANT: James, Judith A.
; APPLICANT: Scofield, R. H.
; TITLE OF INVENTION: PEPTIDE INDUCTION OF AUTOIMMUNITY AND CLINICAL SYMPTOMATOLOG
; NUMBER OF SEQUENCES: 127

US-08-160-604-103
; Sequence 103, Application US/08160604
; Patent No. 6232522
; GENERAL INFORMATION:
; APPLICANT: Harley, John
; APPLICANT: James, Judith A.
; APPLICANT: Scofield, R. H.
; TITLE OF INVENTION: PEPTIDE INDUCTION OF AUTOIMMUNITY AND CLINICAL SYMPTOMATOLOG
; NUMBER OF SEQUENCES: 127

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Patrea L. Pabst
;; STREET: 1100 Peachtree Street, Suite 2800
;; CITY: Atlanta
;; STATE: Georgia
;; COUNTRY: USA
;; ZIP: 30309-4530
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;;
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/160,604
;; FILING DATE: 30-NOV-1993
;; CLASSIFICATION: 424
;;
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/867,819
;; FILING DATE: 13-APR-1992
;;
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/648,205
;; FILING DATE: 31-JAN-1991
;;
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/472,947
;; FILING DATE: 31-JAN-1990
;;
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Pabst, Patrea L.
;; REGISTRATION NUMBER: 31,284
;; REFERENCE/DOCKET NUMBER: OMFLIACIP(3)
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (404)-815-6508
;; TELEFAX: (404)-815-6555
;;
;; INFORMATION FOR SEQ ID NO: 103:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 12 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; HYPOTHEICAL: NO
;; ANTI-SENSE: NO
;; FRAGMENT TYPE: N-terminal
;;
US-08-160-604-103
;;
Query Match 58.9%; Score 33; DB 4; Length 12;
Best Local Similarity 75.0%; Pred. No. 2;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 3 TVEPTGAK 10
Db 2 TVEPTGKR 9
;;
RESULT 8
5185431-9
;; Patent No. 5185431
;; APPLICANT: YOSHIMATSU, KENTARO;SHIKATA, YASUSHI;TANAKA, ISAO;
;; HASEGAWA, YOSHIKAZU;SERO, TOSHIO;OSAWA, TOSHIO
;; TITLE OF INVENTION: RECOMBINANT NATURAL KILLER CELL ACTIVATOR
;; NUMBER OF SEQUENCES: 31
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/07/392,841
;; FILING DATE: 11-AUG-1989
;; SEQ ID NO: 9:
;; LENGTH: 25
5185431-9
;;
Query Match 58.9%; Score 33; DB 6; Length 25;
Best Local Similarity 60.0%; Pred. No. 4.6;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 ADTVEPTGAK 10
;; : : 1 1 1 1 1 1

Db 5 SEFPEPLGAK 14
;;
RESULT 9
US-08-870-518-4
;; Sequence 4, Application US/08870518
;; Patent No. 5925566
;;
;; GENERAL INFORMATION:
;; APPLICANT: Davis, Roger J.
;; APPLICANT: Galcheva-gargova, Zoya
;; TITLE OF INVENTION: NON-ACTIVATED RECEPTOR COMPLEX
;; TITLE OF INVENTION: PROTEINS AND USES THEREOF
;; NUMBER OF SEQUENCES: 35
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Fish & Richardson P.C.
;; STREET: 225 Franklin Street
;; CITY: Boston
;; STATE: MA
;; COUNTRY: US
;; ZIP: 02110-2804
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Diskette
;; COMPUTER: IBM compatible
;; OPERATING SYSTEM: Windows95
;; SOFTWARE: FastSeq for Windows Version 2.0
;;
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/870,518
;; FILING DATE: 06-JUN-1997
;;
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 60/019,219
;; FILING DATE: 06-JUN-1996
;;
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Fasse, Peter J.
;; REGISTRATION NUMBER: 32,983
;; REFERENCE/DOCKET NUMBER: 04020/102001
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 617/542-5070
;; TELEFAX: 617/542-8906
;;
;; INFORMATION FOR SEQ ID NO: 4:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 459 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;;
US-08-870-518-4
;;
Query Match 58.9%; Score 33; DB 2; Length 459;
Best Local Similarity 50.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
QY 1 ADTVEPTGAK 10
Db 80 AETIOPEGTR 89
;;
RESULT 10
US-07-853-913-4
;; Sequence 4, Application US/07853913
;; Patent No. 5338839
;;
;; GENERAL INFORMATION:
;; APPLICANT: McKay, Ronald D.G.
;; APPLICANT: Lendahl, Urban
;; TITLE OF INVENTION: Nestin Expression As An Indicator of
;; TITLE OF INVENTION: Neuroepithelial Tumors
;; NUMBER OF SEQUENCES: 4
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
;; STREET: Two Millitia Drive
;; CITY: Lexington
;; STATE: Massachusetts
;; COUNTRY: U.S.A.
;; ZIP: 02173

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/853,913
FILING DATE: 19920319
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/660,412
FILING DATE: 22-FEB-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/603,803
FILING DATE: 25-OCT-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/201,762
FILING DATE: 02-JUN-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/180,548
ATTORNEY/AGENT INFORMATION:
NAME: Granahan, Patricia
REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: MIT-4641AAAA
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-861-6240
TELEFAX: 617-861-9540
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1618 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-853-913-4

Query Match 58.9%; Score 33; DB 1; Length 1618;
Best Local Similarity 50.0%; Pred. No. 5.1e+02;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 2 DRIVEPTGAK 11
1: 11:11:11
Db 1488 DSAEPGSEE 1497

RESULT 11
US-07-853-913-2
Sequence 2, Application US/07853913
Patent No. 5338839
GENERAL INFORMATION:
APPLICANT: McKay, Ronald D.G.
APPLICANT: Lendahl, Urban
TITLE OF INVENTION: Nestin Expression as An Indicator of
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESS: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Millitia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/853,913
FILING DATE: 19920319
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/660,412
FILING DATE: 22-FEB-1991

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/603,803
FILING DATE: 25-OCT-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/201,762
FILING DATE: 02-JUN-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/180,548
ATTORNEY/AGENT INFORMATION:
NAME: Granahan, Patricia
REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: MIT-4641AAAA
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-861-6240
TELEFAX: 617-861-9540
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1805 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-853-913-2

Query Match 58.9%; Score 33; DB 1; Length 1805;
Best Local Similarity 50.0%; Pred. No. 5.7e+02;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 2 DRIVEPTGAK 11
1: 11:11:11
Db 1679 DSAEPGSEE 1688

RESULT 12
US-08-160-604-104
Sequence 104, Application US/08160604
Patent No. 6232522
GENERAL INFORMATION:
APPLICANT: Harley, John
APPLICANT: James, Judith A.
APPLICANT: Scofield, R. H.
TITLE OF INVENTION: PEPTIDE INDUCTION OF AUTOIMMUNITY AND CLINICAL SYMPTOMATOLOGY
NUMBER OF SEQUENCES: 127
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patrea L. Pabst
STREET: 1100 Peachtree Street, Suite 2800
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30309-4530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/160,604
FILING DATE: 30-NOV-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/867,819
FILING DATE: 13-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/648,205
FILING DATE: 31-JAN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/472,947
FILING DATE: 31-JAN-1990
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: OMRF114C1P(3)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404)-815-6508

TELEFAX: (404)-815-6555
; INFORMATION FOR SEQ ID NO: 104:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
US-08-160-604-104

Query Match 57.1%; Score 32; DB 4; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.9e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 TVEPTG 8
| | | | |
DB 2 TVEPTG 7

RESULT 13
US-09-071-035-300
; Sequence 300, Application US/09071035
; Patent No. 6448043
; GENERAL INFORMATION:
; APPLICANT: Gil H. Choi
; TITLE OF INVENTION: Enterococcus faecalis polynucleotides and polypeptides
; NUMBER OF SEQUENCES: 496
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/071,035
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: A. Anders Brookes
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB369P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 300:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 256 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-071-035-300

Query Match 57.1%; Score 32; DB 4; Length 256;
Best Local Similarity 66.7%; Pred. No. 99;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 TVEPTGAK 11
| | | | |
DB 165 TVEPTDTKD 173

RESULT 14
US-09-071-035-298
; Sequence 298, Application US/09071035
; Patent No. 6448043
; GENERAL INFORMATION:
; APPLICANT: Gil H. Choi
; TITLE OF INVENTION: Enterococcus faecalis polynucleotides and polypeptides
; NUMBER OF SEQUENCES: 496
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/071,035
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: A. Anders Brookes
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB369P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 298:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 284 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-071-035-298

Query Match 57.1%; Score 32; DB 4; Length 284;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 TVEPTGAK 11
| | | | |
DB 192 TVEPTDTKD 200

RESULT 15
US-08-236-918A-2
; Sequence 2, Application US/08236918A
; Patent No. 5674704
; GENERAL INFORMATION:
; APPLICANT: Alderson, Mark R.
; APPLICANT: Goodwin, Raymond G.
; APPLICANT: Smith, Craig A.
; TITLE OF INVENTION: Cytokine Designated 4-1BB Ligand
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kathryn A. Anderson, Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: Washington
; COUNTRY: US
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Power Macintosh

```

; OPERATING SYSTEM: Apple 7.5.3
; SOFTWARE: Microsoft Word, Version #6.0.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/236,918A
; FILING DATE: 06-May-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/060,843
; FILING DATE: 07-May-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Anderson, Kathryn A.
; REGISTRATION NUMBER: 32,172
; REFERENCE/DOCKET NUMBER: 2801-B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 587-0430
; TELEFAX: (206) 233-0644
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 309 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-236-918A-2

```

```

Query Match          57.1%; Score 32; DB 1; Length 309;
Best Local Similarity 66.7%; Pred. No. 1.2e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 ADTVEPTGA 9
   :||| |||
Db 42 SDTVRPTNA 50

```

Search completed: November 18, 2002, 15:56:51
 Job time : 13.7742 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 18, 2002, 15:56:17 ; Search time 6.03226 Seconds
(without alignments)
27.464 Million cell updates/sec

Title: US-09-848-781-3
Perfect score: 56
Sequence: 1 ADVEPTGAK 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 97044 seqs, 15060890 residues

Total number of hits satisfying chosen parameters: 97044

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications_AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
- 12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
- 13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
- 14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	46	82.1	479	10	US-09-971-309-64
2	37	66.1	236	10	US-09-741-669-381
3	37	66.1	459	10	US-09-815-242-11954
4	34	60.7	52	10	US-09-864-761-35593
5	33	58.9	1618	9	US-09-963-875-1
6	32	57.1	422	10	US-09-924-338-2
7	32	57.1	459	9	US-09-997-610-2
8	32	57.1	474	12	US-10-007-693-99
9	32	57.1	577	9	US-09-997-610-6
10	32	57.1	621	10	US-09-953-259-2
11	32	57.1	1237	9	US-10-108-605-211
12	31	55.4	75	10	US-09-864-761-42982
13	31	55.4	353	10	US-09-801-368-116
14	31	55.4	441	10	US-09-924-338-4
15	31	55.4	529	12	US-10-052-586-472
16	31	55.4	548	10	US-09-733-350-4
17	31	55.4	1502	10	US-09-801-368-140
18	30	53.6	27	10	US-09-864-761-33390
19	30	53.6	28	10	US-09-864-761-39493

20	30	53.6	52	10	US-09-925-301-1619	Sequence 1619, Ap
21	30	53.6	124	10	US-09-864-761-37497	Sequence 37497, A
22	30	53.6	142	10	US-09-815-242-11241	Sequence 11241, A
23	30	53.6	223	10	US-09-870-162A-13	Sequence 13, Appl
24	30	53.6	300	10	US-09-764-864-988	Sequence 988, App
25	30	53.6	316	10	US-09-815-242-11929	Sequence 11929, A
26	30	53.6	338	10	US-09-815-242-11750	Sequence 11750, A
27	30	53.6	532	10	US-09-205-658-313	Sequence 313, App
28	30	53.6	560	10	US-09-991-212A-4	Sequence 4, Appl
29	30	53.6	560	10	US-09-915-181A-5	Sequence 5, Appl
30	30	53.6	664	10	US-09-815-242-12577	Sequence 12577, A
31	30	53.6	691	10	US-09-815-242-5223	Sequence 5223, Ap
32	30	53.6	829	9	US-10-053-364-46	Sequence 46, Appl
33	30	53.6	877	10	US-09-746-491-47	Sequence 47, Appl
34	30	53.6	906	10	US-09-905-983-46	Sequence 46, Appl
35	30	53.6	906	10	US-09-746-491-46	Sequence 46, Appl
36	30	53.6	906	10	US-09-746-491-48	Sequence 48, Appl
37	30	53.6	912	10	US-09-905-983-2	Sequence 2, Appl
38	30	53.6	912	10	US-09-746-491-49	Sequence 49, Appl
39	30	53.6	1039	10	US-09-900-237-14	Sequence 14, Appl
40	30	53.6	1537	10	US-09-801-368-104	Sequence 104, Appl
41	30	53.6	1781	10	US-09-738-877-3	Sequence 3, Appl
42	29	51.8	39	10	US-09-864-761-34687	Sequence 34687, A
43	29	51.8	52	10	US-09-864-761-45740	Sequence 45740, A
44	29	51.8	86	9	US-09-747-419-19	Sequence 19, Appl
45	29	51.8	86	9	US-10-035-098-35	Sequence 35, Appl

ALIGNMENTS

RESULT 1
US-09-971-309-64
Sequence 64, Application us/09971309
Patent No. US20020106675A1
GENERAL INFORMATION:
APPLICANT: UEMORI, Takashi
APPLICANT: SATO, Yoshimi
APPLICANT: FUJITA, Tomoko
APPLICANT: MIYAKE, Kazue
APPLICANT: MUKAI, Hiroyuki
APPLICANT: ASADA, Kiyozo
APPLICANT: KATO, Ikunoshin
TITLE OF INVENTION: DNA POLYMERASE-RELATED FACTORS
FILE REFERENCE: 1422-0494P
CURRENT APPLICATION NUMBER: US/09/971,309
CURRENT FILING DATE: 2001-10-05
PRIOR APPLICATION NUMBER: US 09/446,504
PRIOR FILING DATE: 1999-12-23
PRIOR APPLICATION NUMBER: PCT/JP98/02845
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: JP 9-187496
PRIOR FILING DATE: 1997-06-26
PRIOR APPLICATION NUMBER: JP 9-320692
PRIOR FILING DATE: 1997-11-27
NUMBER OF SEQ ID NOS: 92
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 64
LENGTH: 479
TYPE: PRT
ORGANISM: Pyrococcus furiosus
US-09-971-309-64

Query Match Best local Similarity 82.1%; Score 46; DB 10; Length 479;

Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 ADVEPTGAK 11
||:|:|:|:|
DB 110 ADNIEPTGAK 120

RESULT 2

```

US-09-741-669-381
; Sequence 381, Application US/09741669
; Patent No. US20020022718A1
; GENERAL INFORMATION:
; APPLICANT: Forsyth, R. Allyn
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; TITLE OF INVENTION: Genes identified as required for
; FILE REFERENCE: ELITRA.009A
; CURRENT APPLICATION NUMBER: US/09/741,669
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 60/173005
; NUMBER OF SEQ ID NOS: 481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 381
; LENGTH: 236
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-741-669-381

```

```

Query Match      66.1%; Score 37; DB 10; Length 236;
Best Local Similarity 80.0%; Pred. No. 3.7;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

Qy      1 ADTVEPTGAK 10
      |||||
Db      69 ADLVEATGAK 78

```

```

RESULT 3
US-09-815-242-11954
; Sequence 11954, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11954
; LENGTH: 459
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-815-242-11954

```

```

Query Match      66.1%; Score 37; DB 10; Length 459;
Best Local Similarity 87.5%; Pred. No. 7.9;

```

```

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy      1 ADTVEPTG 8
      |||||
Db      105 ADTVEPTG 112

```

```

RESULT 4
US-09-864-761-35593
; Sequence 35593, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Aecmlca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
; SEQ ID NO 35593
; LENGTH: 52
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC008080.1
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.3
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.9

```

OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.6
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.6
OTHER INFORMATION: EST_HUMAN HIT: AA230303.1, EVALUATE 1.00e-18
OTHER INFORMATION: SWISSPROT HIT: P19158, EVALUATE 4.00e-01
US-09-864-761-35593

Query Match 60.7%; Score 34; DB 10; Length 52;
Best Local Similarity 54.5%; Pred. No. 2.7;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 ADVEPTGAK 11
DB 14 SFTIRVSGAKE 24

RESULT 5
US-09-963-875-1
Sequence 1, Application US/09963875
Patent No. US20020164307A1
GENERAL INFORMATION:

APPLICANT: Massachusetts General Hospital
TITLE OF INVENTION: Stem Cells of the Islets of Langerhans and Their Use in Treating
TITLE OF INVENTION: Mellitus
FILE REFERENCE: 17633/1235

CURRENT APPLICATION NUMBER: US/09/963,875

CURRENT FILING DATE: 2001-09-26

PRIOR APPLICATION NUMBER: US60/169082

PRIOR FILING DATE: 1999-12-06

PRIOR APPLICATION NUMBER: US 60/215109

PRIOR FILING DATE: 2000-06-28

PRIOR APPLICATION NUMBER: US 60/238880

PRIOR FILING DATE: 2000-10-06

PRIOR APPLICATION NUMBER: US 09/731261

PRIOR FILING DATE: 2000-12-06

NUMBER OF SEQ ID NOS: 58

SOFTWARE: PatentIn version 3.1

SEQ ID NO 1

LENGTH: 1618

TYPE: PRT

ORGANISM: Homo sapiens

US-09-963-875-1

Query Match 58.9%; Score 33; DB 9; Length 1618;
Best Local Similarity 50.0%; Pred. No. 2e+02;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 2 DIVEPTGAK 11
DB 1488 DSAEPGSEE 1497

RESULT 6
US-09-924-338-2
Sequence 2, Application US/09924338
Patent No. US20020082233A1

GENERAL INFORMATION:

APPLICANT: Tobin, James

TITLE OF INVENTION: HUMAN INTERLEUKIN-11 RECEPTOR

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Genetics Institute, Inc.

STREET: 87 Cambridgepark Drive

CITY: Cambridge

STATE: MA

COUNTRY: USA

ZIP: 02140

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

TITLE OF INVENTION: AND DIAGNOSIS OF CHLAMYDIAL INFECTION

FILE REFERENCE: 210121.515C2

FILING DATE: 07-Aug-2001
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/151,102
FILING DATE: 1998-09-10
APPLICATION NUMBER: 08/362,304
FILING DATE: 22-DEC-1994

ATTORNEY/AGENT INFORMATION:

NAME: Brown, Scott A.

REGISTRATION NUMBER: 32,724

REFERENCE/DOCKET NUMBER: G15252

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 498-8224

TELEFAX: (617) 876-5851

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 422 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-09-924-338-2

Query Match 57.1%; Score 32; DB 10; Length 422;
Best Local Similarity 66.7%; Pred. No. 69;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 TVEPTGAK 11
DB 270 TVEPTGAK 278

RESULT 7
US-09-997-610-2
Sequence 2, Application US/09997610
Patent No. US20020156244A1

GENERAL INFORMATION:

APPLICANT: Fox, Brian

TITLE OF INVENTION: ADIPOCYTE COMPLEMENT RELATED PROTEIN

TITLE OF INVENTION: ZACRP13

FILE REFERENCE: 00-96

CURRENT APPLICATION NUMBER: US/09/997,610

CURRENT FILING DATE: 2001-11-29

PRIOR APPLICATION NUMBER: US 60/253,924

PRIOR FILING DATE: 2000-11-29

NUMBER OF SEQ ID NOS: 7

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 2

LENGTH: 459

TYPE: PRT

ORGANISM: Homo sapiens

US-09-997-610-2

Query Match 57.1%; Score 32; DB 9; Length 459;
Best Local Similarity 75.0%; Pred. No. 76;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 VEPTGAK 11
DB 271 VEPTGAK 278

RESULT 8
US-10-007-693-99
Sequence 99, Application US/10007693
Patent No. US20020146776A1

GENERAL INFORMATION:

APPLICANT: Bhatia, Ajay

TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT

TITLE OF INVENTION: AND DIAGNOSIS OF CHLAMYDIAL INFECTION

FILE REFERENCE: 210121.515C2

; CURRENT APPLICATION NUMBER: US/10/007,693
; CURRENT FILING DATE: 2001-12-05
; NUMBER OF SEQ ID NOS: 157
; SEQ ID NO 99
; LENGTH: 474
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis serovar D
US-10-007-693-99

Query Match 57.1%; Score 32; DB 12; Length 474;
Best Local Similarity 77.8%; Pred. No. 78;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 3 TVEPTGAK 11
||| |||
DB 209 TVSPYTAKE 217

RESULT 9
US-09-997-610-6
; Sequence 6, Application US/09997610
; Patent No. US20020156244A1
; GENERAL INFORMATION:
; APPLICANT: Fox, Brian
; TITLE OF INVENTION: HOLLOWAY, James L.
; TITLE OF INVENTION: ADIPOCYTE COMPLIMENT RELATED PROTEIN
; FILE REFERENCE: 00-96
; CURRENT APPLICATION NUMBER: US/09/997,610
; CURRENT FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/253,924
; PRIOR FILING DATE: 2000-11-29
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 577
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-997-610-6

Query Match 57.1%; Score 32; DB 9; Length 577;
Best Local Similarity 75.0%; Pred. No. 98;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 4 VEPGAK 11
||| |||
DB 389 VEPGAKK 396

RESULT 10
US-09-953-259-2
; Sequence 2, Application US/09953259
; Patent No. US20020115159A1
; GENERAL INFORMATION:
; APPLICANT: FARMICK, Mike
; APPLICANT: HUTTMACHER, Klaus
; APPLICANT: PFEFFERLE, Walter
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES CODING FOR THE ATR61 PROTEIN
; FILE REFERENCE: 213903US0X
; CURRENT APPLICATION NUMBER: US/09/953,259
; CURRENT FILING DATE: 2001-09-17
; PRIOR APPLICATION NUMBER: DE10045579.4
; PRIOR FILING DATE: 2000-09-15
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 621
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-953-259-2

Query Match 57.1%; Score 32; DB 10; Length 621;
Best Local Similarity 63.6%; Pred. No. 11e+02;

Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
OY 1 ADVEPTGAK 11
| |||| |
DB 322 AQOVEPTANKE 332

RESULT 11
US-10-108-605-211
; Sequence 211, Application US/10108605
; Patent No. US20020160934A1
; GENERAL INFORMATION:
; APPLICANT: Broadus, Julie
; APPLICANT: Stam, Lynn
; APPLICANT: Bachmann, Jane
; APPLICANT: Kandari, Kim
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCES FROM DROSOPHILA MELANOGASTER THAT ENCODE
; TITLE OF INVENTION: PROTEINS ESSENTIAL FOR LARVAL VIABILITY AND USES THEREOF
; FILE REFERENCE: 31133B
; CURRENT APPLICATION NUMBER: US/10/108,605
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: US 09/761,142
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/176,418
; PRIOR FILING DATE: 2000-01-14
; NUMBER OF SEQ ID NOS: 361
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 211
; LENGTH: 1237
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-10-108-605-211

Query Match 57.1%; Score 32; DB 9; Length 1237;
Best Local Similarity 85.7%; Pred. No. 2.3e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 DVEPTG 8
||| |||
DB 92 DVEPTG 98

RESULT 12
US-09-864-761-42982
; Sequence 42982, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecmca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669

Sequence 472, Application US/10052586
Patent No. US20020127584A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C1
CURRENT APPLICATION NUMBER: US/10/052,586
CURRENT FILING DATE: 2002-01-15
PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059266
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/063120
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063121
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063486
PRIOR FILING DATE: 1997-10-21
PRIOR APPLICATION NUMBER: 60/063540
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063541
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063544
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063564
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063734
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063870
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/064103
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066120
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/066466
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/066772
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/069335
PRIOR FILING DATE: 1997-12-11
PRIOR APPLICATION NUMBER: 60/069425
PRIOR FILING DATE: 1997-12-12
PRIOR APPLICATION NUMBER: 60/069870
PRIOR FILING DATE: 1997-12-17
PRIOR APPLICATION NUMBER: 60/068017
PRIOR FILING DATE: 1997-12-18
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27

PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081070
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081195
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081838
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083559
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/084366
PRIOR FILING DATE: 1998-05-05
PRIOR APPLICATION NUMBER: 60/084414
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084639
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084640
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084643
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/085573
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085580
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085582
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085700
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/086023
PRIOR FILING DATE: 1998-05-18
PRIOR APPLICATION NUMBER: 60/086392
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/086486
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/087098
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087208
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088025

;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088028
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088029
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088033
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088167
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088202
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088212
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088217
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088326
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088655
;; PRIOR FILING DATE: 1998-06-09
;; PRIOR APPLICATION NUMBER: 60/088722
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088738
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088740
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088811
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088824
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088825
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088826
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088861
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/088863
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/088876
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/089090
;; PRIOR FILING DATE: 1998-06-12
;; PRIOR APPLICATION NUMBER: 60/089105
;; PRIOR FILING DATE: 1998-06-12
;; PRIOR APPLICATION NUMBER: 60/089512
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089514
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089538
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089598
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089653
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089908

Query Match 55.4%; Score 31; DB 12; Length 529;

Best Local Similarity 75.0%; Pred. No. 1.4e+02;

Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 VEPTGAKE 11
111111
DB 514 VEPVGEKE 521

Search completed: November 18, 2002, 16:05:45
Job time : 6.03226 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 18, 2002, 15:53:18 ; Search time 13.129 Seconds
(without alignments)
80.545 Million cell updates/sec

Title: US-09-848-781-3

Perfect score: 56

Sequence: 1 ADTVEPTGAK 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: pIR1:*
2: pIR2:*
3: pIR3:*
4: pIR4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	44	78.6	289	2 S20888	phospholipase C (E
2	44	78.6	289	2 C43868	lecithinase - list
3	44	78.6	289	2 AF1100	phospholipase C (l
4	40	71.4	468	2 G71231	probable replicati
5	40	71.4	479	2 D75198	activator 1, repli
6	37	66.1	211	2 AF0644	probable TetR-fam
7	37	66.1	219	2 AC2051	hypothetical prote
8	37	66.1	236	2 A90815	hypothetical prote
9	37	66.1	236	2 E85674	hypothetical prote
10	37	66.1	236	2 D64855	probable transcrip
11	37	66.1	459	2 G83174	exodeoxyribonuclea
12	37	66.1	1819	2 T26533	hypothetical prote
13	36	64.3	159	2 C84299	hypothetical prote
14	36	64.3	252	2 AC0391	conserved hypotet
15	36	64.3	534	2 D98224	dipeptide transpor
16	36	64.3	534	2 AC3062	hypothetical prote
17	36	64.3	512	2 T15485	hypothetical prote
18	36	64.3	642	2 S77012	cadmium-transporti
19	36	64.3	1892	2 T18314	hypothetical prote
20	35	62.5	234	2 S36348	hypothetical prote
21	35	62.5	260	2 C83618	probable CCA trans
22	35	62.5	261	2 S16610	opacty protein op
23	35	62.5	304	2 D82304	glutamyl-tRNA synt
24	35	62.5	338	2 S16613	opacty protein op
25	35	62.5	1025	2 H86250	hypothetical prote
26	35	62.5	1224	2 T19749	hypothetical prote
27	35	62.5	1484	2 T29275	hypothetical prote
28	34	60.7	65	2 E75022	ssu ribosomal prot
29	34	60.7	204	2 AF0053	osmotically induci

30	34	60.7	223	2 S75567	hypothetical prote
31	34	60.7	254	2 T23323	hypothetical prote
32	34	60.7	452	2 AE1710	hypothetical prote
33	34	60.7	452	2 AG1339	hypothetical prote
34	34	60.7	452	2 T37704	zinc-finger protei
35	34	60.7	534	2 C70548	probable regulator
36	34	60.7	779	2 E83157	hypothetical prote
37	34	60.7	880	2 S03601	RNA-binding protei
38	34	60.7	881	1 P2XRSR	RNA-binding protei
39	34	60.7	881	1 P2XRUX	RNA-binding protei
40	34	60.7	890	1 P2XRWA	RNA-binding protei
41	34	60.7	1516	2 F83085	conserved hypotet
42	34	60.7	2171	2 E86342	hypothetical prote
43	33.5	59.8	2508	2 S61441	surface-associated
44	33	58.9	86	2 T46106	hypothetical prote
45	33	58.9	138	2 F81243	Dnak suppressor pr

ALIGNMENTS

RESULT 1

S20888 phospholipase C (EC 3.1.4.3) precursor - Listeria monocytogenes (fragment)

C/Species: Listeria monocytogenes

C/Date: 16-Sep-1992 #sequence_rev1510n 16-Sep-1992 #text_change 24-Sep-1999

C/Accession: S20888

R/Domann, E.; Wehland, J.; Rohde, M.; Pistor, S.; Hartl, M.; Goebel, W.; Leimeister-W

EMBO J. 11, 1981-1990, 1992

A/Title: A novel bacterial virulence gene in Listeria monocytogenes required for host

A/Reference number: S20887; M0ID:92258410; PMID:1582425

A/Accession: S20888

A/Molecule type: DNA

A/Residues: 1-289 <DOM>

A/Cross-references: EMBL:X59723; NID:g44091; PIDN:CAA42408.1; PTD:g44093

C/Genetics:

A/Gene: pICB

C/Superfamily: Bacillus cereus phospholipase C

C/Keywords: phosphoric diester hydrolase

F:1-25/Domain: signal sequence #status predicted <Sig>

F:26-289/Product: phospholipase C #status predicted <Mat>

Query Match Best Local Similarity 78.6%; Score 44; DB 2; Length 289;

Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 DTVEPTGAK 10
|||||||

DB 260 DTVEPTGAR 268

RESULT 2

C43868 lecithinase - Listeria monocytogenes

C/Species: Listeria monocytogenes

C/Date: 10-Mar-1993 #sequence_rev1510n 18-Nov-1994 #text_change 24-Sep-1998

C/Accession: C43868

R/Vazquez-Boland, J. A.; Kocks, C.; Dramai, S.; Ohayon, H.; Geoffroy, C.; Mengaud, J.;

Infect. Immun. 60, 219-230, 1992

A/Title: Nucleotide sequence of the lecithinase operon of Listeria monocytogenes and

A/Reference number: A43868; M0ID:92104678; PMID:1309513

A/Accession: C43868

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-289 <VAZ>

A/Note: sequence inconsistent with the nucleotide translation

A/Note: sequence extracted from NCBI backbone (NCBIN:74437, NCBI:74462)

C/Superfamily: Bacillus cereus phospholipase C

Query Match Best Local Similarity 78.6%; Score 44; DB 2; Length 289;

Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 DIVEPTGAK 10
|||||||
Db 260 DIVEPTGAR 268

RESULT 3

AF1100
phospholipase C [imported] - *Listeria monocytogenes* (strain EGD-e)
C:Species: *Listeria monocytogenes*
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C:Accession: AF1100
R:Jaaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihl, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurpkat, G.; Madueno, E.; Maltournam, A.; Ma
ok, C.; Schleuter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend,
A.; Title: Comparative genomics of *Listeria* species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AF1100
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-289 <GLA>
A:Cross-references: GB:NC_003210; PIDN:CAD00732.1; PID:g16409570; GSPDB:GN00177
A:Experimental source: strain EGD-e
C:Genetics:
A:Gene: plcB
C:Superfamily: *Bacillus cereus* phospholipase C

Query Match 78.6%; Score 44; DB 2; Length 289;
Best Local Similarity 88.9%; Pred. No. 0.69;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 DIVEPTGAK 10
|||||||
Db 260 DIVEPTGAR 268

RESULT 4

G71231
probable replication factor C subunit - *Pyrococcus horikoshii*
C:Species: *Pyrococcus horikoshii*
C:Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 21-Jul-2000
C:Accession: G71231
R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Hakiwara, Y.; Hino, Y.; Yamamoto, S.; Seki
M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kishida, N.; Oguchi
DNA Res. 5, 55-76, 1998
A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic
A:Reference number: A71000; MUID:98344137; PMID:9679194
A:Accession: G71231
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-468 <KAM>
A:Cross-references: GB:AP000001; NID:g3236128; PIDN:BA029182.1; PID:g3256499
A:Experimental source: strain OT3
A:Note: this accession replaces an interim accession for a sequence replaced by GenBank
C:Genetics:
A:Gene: PH0113
C:Superfamily: replication factor C large chain

Query Match 71.4%; Score 40; DB 2; Length 468;
Best Local Similarity 63.6%; Pred. No. 7.1;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 ADIVEPTGAK 11
||:|:|:|
Db 110 ADNIERSGAP 120

RESULT 5

D75198
activator 1, replication factor C, large s ubunit PAB0069 - *Pyrococcus abyssi* (strain Or
C:Species: *Pyrococcus abyssi*

C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000
C:Accession: D75198
R:anonymous, Genoscope
submitted to the EMBL Data Library, July 1999
A:Description: *Pyrococcus abyssi* genome sequence: insights into archaeal chromosome s
A:Reference number: A75001
A:Accession: D75198
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-479 <KAM>
A:Cross-references: GB:AJ248283; GB:AL096836; NID:g5457433; PIDN:CA849035.1; PID:g545
A:Experimental source: strain Orsay
C:Genetics:
A:Gene: rfcL; PAB0069
C:Superfamily: replication factor C large chain

Query Match 71.4%; Score 40; DB 2; Length 479;
Best Local Similarity 63.6%; Pred. No. 7.2;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 ADIVEPTGAK 11
||:|:|:|
Db 110 ADNIERSGAP 120

RESULT 6

AF0644
probable TetR-family regulatory protein STY1253 [imported] - *Salmonella enterica* subs
C:Species: *Salmonella enterica* subsp. *enterica* serovar Typh
A:Note: this species has also been called *Salmonella typh*
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 09-Nov-2001
C:Accession: AF0644
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Church
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farr
, S.; Mout, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens,
A.; Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* se
A:Reference number: AB0502; PMID:11677608
A:Accession: AF0644
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-211 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD08337.1; PID:g16502382; GSPDB:GN00176
C:Genetics:
A:Gene: STY1253

Query Match 66.1%; Score 37; DB 2; Length 211;
Best Local Similarity 80.0%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ADIVEPTGAK 10
||:|:|:|
Db 43 ADIVEATGAR 52

RESULT 7

AC2051
hypothetical protein alr1961 [imported] - *Nostoc* sp. (strain PCC 7120)
C:Species: *Nostoc* sp.
A:Note: *Nostoc* sp. strain PCC 7120 is a synonym of *Anabaena* sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002
C:Accession: AC2051
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kurita, T.; Sasamoto, S.; Matsumae, A.; Irigu
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AC2051
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-219 <KUR>
A:Cross-references: GB:BA000019; PIDN:BA073660.1; PID:g17131051; GSPDB:GN00179

A:Experimental source: strain FCC 7120
C:Genetics:
A:Gene: alr1961

Query Match
Best Local Similarity 66.1%; Score 37; DB 2; Length 219;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 ADTVEPTGAK 10
||| |||||
Db 128 ADLVEATGAK 137

RESULT 8
A:Gene: Ecs1489
A:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
C:Accession: A90815
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.; Sasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genomic reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: A90815
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-236 <HAY>
A:Cross-references: GB:BA000007; PIDN:BAH34912.1; PID:q13360953; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain RIMD 0509952
C:Genetics:
A:Gene: Ecs1489

Query Match
Best Local Similarity 66.1%; Score 37; DB 2; Length 236;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 ADTVEPTGAK 10
||| |||||
Db 69 ADLVEATGAK 78

RESULT 9
A:Gene: E85674
A:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C:Accession: E85674
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca, Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11205551
A:Accession: E85674
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-236 <STO>
A:Cross-references: GB:AE005174; NID:q12514658; PIDN:AA655857.1; GSPDB:GN00145; UMG:217
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: ycfQ

Query Match
Best Local Similarity 66.1%; Score 37; DB 2; Length 236;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 ADTVEPTGAK 10
||| |||||
Db 69 ADLVEATGAK 78

RESULT 10
D64855

probable transcription regulator ycfQ - Escherichia coli (strain K-12)
C:Species: Escherichia coli
C:Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 01-Mar-2002
C:Accession: D64855
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: D64855
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-236 <BLAT>
A:Cross-references: GB:AE00211; GB:U00096; NID:q1787345; PIDN:AACT4195.1; PID:q17873
A:Experimental source: strain K-12, substrain MG1655
C:Genetics:
A:Gene: ycfQ
C:Keywords: DNA binding; transcription regulation
F:57-86/Region: helix-turn-helix motif

Query Match
Best Local Similarity 66.1%; Score 37; DB 2; Length 236;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 ADTVEPTGAK 10
||| |||||
Db 69 ADLVEATGAK 78

RESULT 11
A:Gene: G83174
A:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: G83174
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Adam, S.; Van, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen.
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: G83174
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-459 <STO>
A:Cross-references: GB:AE004796; GB:AE004091; NID:q9949939; PIDN:AA607164.1; GSPDB:GN
A:Experimental source: strain PA01
C:Genetics:
A:Gene: xseA; PA3777
C:Superfamily: exodeoxyribonuclease VII

Query Match
Best Local Similarity 66.1%; Score 37; DB 2; Length 459;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 ADTVEPTG 8
||| |||||
Db 105 ADLVEATG 112

RESULT 12
A:Gene: T26533
A:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T26533
R:Harris, B.
Submitted to the EMBL Data Library, December 1998
A:Reference number: Z20226
A:Accession: T26533
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1819 <WIL>

A:Cross-references: EMBL:AL034393; PIDN:CAA22324.1; CESP:Y18D10A.13
A:Experimental source: clone Y18D10A
C:Genetics:
A:Gene: CESP:Y18D10A.13
A:introns: 51/3; 83/1; 169/2; 232/3; 310/1; 469/1; 528/3; 599/1; 619/1; 949/1; 1087/2; 1

Query Match 66.1%; Score 37; DB 2; Length 1819;
Best Local Similarity 87.5%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 DVEPTGA 9
DB 1077 DVEPTGA 1084

RESULT 13
C84299
hypothetical protein Vng1453h [imported] - Halobacterium sp. NRC-1
C:Species: Halobacterium sp. NRC-1
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: C84299
R:Ng, W.V.; Kennedy, S.P.; Mahafas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Laaky, S.; Leihanser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablo Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.; Li A:Title: Genome sequence of Halobacterium species NRC-1.
A:Reference number: A84160; MUID:20504483; PMID:11016950
A:Accession: C84299
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-159 <STO>
A:Cross-references: GB:AE004437; NID:g10580949; PIDN:AGI9759.1; GSPDB:GN00138
C:Genetics:
A:Gene: VNG1453H

Query Match 64.3%; Score 36; DB 2; Length 159;
Best Local Similarity 60.0%; Pred. No. 14;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 ADVEPTGAK 10
DB 141 ADSEPTGPR 150

RESULT 14
AC0391
conserved hypothetical protein YPO3220 [imported] - Yersinia pestis (strain CO92)
C:Species: Yersinia pestis
C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 02-Nov-2001
C:Accession: AC0391
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titchell, R.W.; Holden, M.T.G.; Prentice, M.B. deno-Parraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; Il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell, Nature 413, 523-527, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: AC0391
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-252 <KUR>
A:Cross-references: GB:AL590842; PIDN:CAC92455.1; PID:g15981156; GSPDB:GN00175
C:Genetics:
A:Gene: YPO3220

Query Match 64.3%; Score 36; DB 2; Length 252;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 VEPTGAK 10
DB 81 VEPTGAK 87

RESULT 15
D98224
dipeptide transport protein (AB036425) [imported] - Agrobacterium tumefaciens (strain C:Species: Agrobacterium tumefaciens
C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 11-Jan-2002
C:Accession: D98224
R:Goodner, B.; Hinkle, G.; Galtung, S.; Miller, N.; Blanchard, M.; Quorllo, B.; Gold A.; Liu, F.; Wollam, C.; Allinger, M.; Doughy, D.; Scott, C.; Lappas, C.; Markelz, Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium A:Reference number: A97359; PMID:11743194
A:Accession: D98224
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-534 <KUR>
A:Cross-references: GB:AE007870; PIDN:AAK89318.1; PID:g15159156; GSPDB:GN00170
C:Genetics:
A:Gene: AGR_L1481
A:Map position: linear chromosome
C:Superfamily: dipeptide transport protein

Query Match 64.3%; Score 36; DB 2; Length 534;
Best Local Similarity 63.6%; Pred. No. 50;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 ADVEPTGAK 11
DB 188 ADLEKAGTKE 198

Search completed: November 18, 2002, 15:56:10
Job time : 14.129 secs


```

RT phospholipase C from Listeria monocytogenes.;
RL Infect. Immun. 59:2382-2388(1991).
CC -1- FUNCTION: IMPORTANT ROLE IN THE INFECTIOUS PROCESS. MAY
CC CONTRIBUTE TO EFFICIENT LYSIS OF THE TWO-MEMBRANE VACUOLES THAT
CC SURROUND THE BACTERIA AFTER DIRECT CELL-TO-CELL SPREAD.
CC -1- CATALYTIC ACTIVITY: A phosphatidylcholine + H(2)O = 1,2-
CC diacylglycerol + choline phosphate.
CC -1- COFACTOR: BINDS 3 ZINC IONS (By similarity).
CC -1- SIMILARITY: STRONG, TO B.CEREUS PHOSPHOLIPASE C.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M62881; AAA25270.1; -
DR EMBL: X59723; CAA42408.1; -
DR EMBL: AL591974; CAD00732.1; -
DR PIR: C43868; C43868.
DR HSSP: P09598; 1AH7.
DR Listlist: LMO00205; -
DR InterPro: IPR001531; Zn_dep_PLPC.
DR Pfam: PF00882; Zn_dep_PLPC; 1.
DR PRINTS: PR00479; PRPHPLPASEC.
DR ProDom: PD003946; Zn_dep_PLPC; 1.
DR PROSITE: PS00384; PROKAR_ZN_DEPEND_PLPC; 1.
DR K1M Hydroxylase; Zinc; Zymogen; Signal; Complete proteome.
FT SIGNAL 1 25
FT PROPEP 26 51
FT CHAIN 52 289
FT METAL 52 289
FT METAL 52 289
FT METAL 65 65
FT METAL 106 106
FT METAL 120 120
FT METAL 169 169
FT METAL 173 173
FT METAL 179 179
FT METAL 193 193
FT METAL 197 197
FT CONFLICT 5 5
FT CONFLICT 13 13
FT CONFLICT 222 222
SQ SEQUENCE 289 AA; 33277 MW; 0F35A2A3EDA6E372 CRC64;

Query Match 78.6%; Score 44; DB 1; Length 289;
Best Local Similarity 88.9%; Pred. No. 0.28;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 DIVEPTGAK 10
Db 260 DIVEPTGAR 268
|||||:

```

```

RA Blatner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1453-1474(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-K12;
RX MEDLINE=97061202; PubMed=8905232;
RA Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,
RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
RA Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,
RA Sempel G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
RA Yano M., Horinouchi T.;
RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 12.7-28.0 min region on the linkage map.";
RL DNA Res. 3:137-155(1996).
CC -1- SIMILARITY: BELONGS TO THE TETR/ACR FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE000211; AAC74195.1; ALT_INIT.
DR EMBL: D90746; BAA35926.1; ALT_INIT.
DR EMBL: D90747; BAA35931.1; ALT_INIT.
DR Ecogene: EG13435; Ycfo.
DR InterPro: IPR001647; HTH_Tetr.
DR Pfam: PF00440; Tetr; 1.
DR PROSITE: PS01081; HTH_TETR_FAMILY; FALSE_NEG.
DR Hypothetical protein; Transcription regulation; DNA-binding;
KW Complete proteome.
FT DNA_BIND 41 60
FT SEQUENCE 210 AA; 23361 MW; 8C1D525668756842 CRC64;

Query Match 66.1%; Score 37; DB 1; Length 210;
Best Local Similarity 80.0%; Pred. No. 4.8;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 ADIVEPTGAK 10
Db 43 ADIVEATGAK 52
|||||:

```

RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.,
RT "Complete genome sequence of *Pseudomonas aeruginosa* PAO1, an
RT opportunistic pathogen." ;
RL Nature 406:959-964(2000).
CC -1- FUNCTION: BIDIRECTIONALLY DEGRADABLE SINGLE-STRANDED DNA INTO LARGE
CC ACID-INSOLUBLE OLIGONUCLEOTIDES, WHICH ARE THEN DEGRADED FURTHER
CC INTO SMALL ACID-SOLUBLE OLIGONUCLEOTIDES (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: Exonuclease activity cleavage in either 5' -
CC or 3' - to 5' -direction to yield nucleoside 5'-phosphates.
CC -1- SUBUNIT: HETEROOLIGOMER COMPOSED OF LARGE AND SMALL SUBUNITS (BY
CC SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE XSEPA FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, AE004796; AAC07164.1. -
DR InterPro: IPR003753; Exonuc-VII.L.
DR InterPro: IPR004365; tRNA_antl.
DR Pfam: PF01336; tRNA_antl. 1.
DR Pfam: PF02601; Exonuc-VII.L. 1.
DR TIGRfam: TIGR00237; xseA. 1.
KW Hydrolyase; Nuclease; Exonuclease; Complete proteome.
SQ SEQUENCE 459 AA: 51231 MW; CQFAFB85D4174935 CRC64;

Query Match 66.1%; Score 37; DB 1; Length 459;
Best Local Similarity 87.5%; Pred. No. 11;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ADVEPTG 8
| | | | | | |
DB 105 ADVEPAG 112

RESULT 4
OP65_NEIGO STANDARD; PRT; 234 AA.
ID OP65_NEIGO
AC 004885;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Opacity protein OPA65 precursor (Fragment).
OS *Neisseria gonorrhoeae*.
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; *Neisseria*.
OX NCBI_TaxID=485;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VPI;
RX MEDLINE=93178439; PubMed=8440254;
RA Kusch E.-M., Knepper B., Kuroki T., Heuer I., Meyer T.F.;
RT "Variable opacity (opa) outer membrane proteins account for the cell
RT tropisms displayed by *Neisseria gonorrhoeae* for human leukocytes and
RT epithelial cells." ;
RL EMBL J.12:641-650(1993).
CC -1- FUNCTION: IMPLICATED IN A NUMBER OF ADHERENCE FUNCTIONS. OPA
CC PROTEINS ARE IMPLICATED IN PATHOGENESIS AND ARE SUBJECT TO PHASE
CC VARIATION.
CC -1- SUBCELLULAR LOCATION: Outer membrane.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

CC EMBL, Z18940; CAA79373.1; -.
DR PIR: S28617; S28617.
DR InterPro: IPR003394; Porin_opacity.
DR Pfam: PF02462; Opacity. 1.
KW Outer membrane; Multigene family; Signal.
FT NON_TER 1
FT SIGNAL 1
FT CHAIN 2
FT NON_TER 234
FT NON_TER 234
SQ SEQUENCE 234 AA: 26242 MW; D66A0BA8424C2F1 CRC64;

Query Match 62.5%; Score 35; DB 1; Length 234;
Best Local Similarity 70.0%; Pred. No. 14;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 DVEPTGAK 11
| | | | | | |
DB 25 DVEPTGAKK 34

RESULT 5
YC70_CORGL
ID YC70_CORGL STANDARD; PRT; 491 AA.
AC P42531;
DT 01-NOV-1995 (Rel. 32, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein Cg11270.
GN Cg11270.
OS *Corynebacterium glutamicum* (Brevibacterium flavum).
OC Bacteria; Actinobacteria; Actinobacteriales; Actinobacteridae;
OC Actinomycetales; Corynebacteriaceae; Corynebacteriaceae;
OC Corynebacterium.
OX NCBI_TaxID=1718;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RA Nakagawa S.;
RT "Complete genomic sequence of *Corynebacterium glutamicum* ATCC 13032." ;
RT Submitted (MAY-2002) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE OF 1-62 FROM N.A.
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RX MEDLINE=93374855; PubMed=8366043;
RA Kelhauer C., Eggeling L., Sahn H.;
RT "Isolucine synthesis in *Corynebacterium glutamicum*: molecular
RT analysis of the *ilvB-ilvN-ilvC* operon." ;
RL J. Bacteriol. 175:5595-5603(1993).
CC -1- SUBCELLULAR LOCATION: Membrane-associated (Potential).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL, AP005278; BAB98663.1; -.
DR EMBL, L09232; AAA62428.1; -.
KW Hypothetical protein; Transmembrane.
FT TRANSMEM 21
FT POTENTIAL 43
SQ SEQUENCE 491 AA: 52551 MW; 599EDD5F60CAED4 CRC64;

Query Match 62.5%; Score 35; DB 1; Length 491;
Best Local Similarity 85.7%; Pred. No. 30;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ADVEPT 7
| | | | | | |
DB 450 ADVEPT 456

RESULT 6
RS27_PYPFU STANDARD; PRT; 63 AA.
ID Q8U474;
AC 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DR 15-JUN-2002 (Rel. 41, Last annotation update)
DE 30S ribosomal protein S27e.
GN RPS27E OR PF0218.
OS Pyrococcus furiosus.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
CC Pyrococcus.
NCBI_TaxID=2261;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-Vcl / DSM 3638 / ATCC 43587 / JCM 8422;
RA Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.;
RT "The complete sequence of the Pyrococcus furiosus genome."
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
CC -1- MISCELLANEOUS: BINDS ONE ZINC ION PER MOLECULE (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE S27E FAMILY OF RIBOSOMAL PROTEINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
DR EMBL: AE010148; AL080342.1;
DR PROSITE: PS01168; RIBOSOMAL_S27E; 1.
KW Ribosomal protein; Zinc-finger; Metal-binding; Complete proteome.
FT ZN_FING 18 CA-TYPE
SQ SEQUENCE 63 AA; 6877 MW; 51AC0CB2E9988A0E CRC64;
Query Match 60.7%; Score 34; DB 1; Length 63;
Best Local Similarity 70.0%; Pred. No. 5.2;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
OY 1 ADIPEPTGAK 10
DB 42 ATLVEPTGK 51
RESULT 7
RS27_PYPAB STANDARD; PRT; 65 AA.
ID Q9UX23;
AC 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DR 16-OCT-2001 (Rel. 40, Last annotation update)
DE 30S ribosomal protein S27e.
GN RPS27E OR PAB7435.
OS Pyrococcus abyssi.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
CC Pyrococcus.
NCBI_TaxID=29292;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-GE5 / Orsay;
RA Hellig R.;
RT "Pyrococcus abyssi genome sequence: insights into archaeal chromosome
structure and evolution."
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
CC -1- MISCELLANEOUS: BINDS ONE ZINC ION PER MOLECULE (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE S27E FAMILY OF RIBOSOMAL PROTEINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
DR EMBL: AJ246288; CAB50619.1;
DR InterPro: IPR000592; Ribosomal_S27E.
DR Pfam: PF01667; Ribosomal_S27e; 1.
DR ProDom: PD004466; Ribosomal_S27E; 1.
DR PROSITE: PS01168; RIBOSOMAL_S27E; 1.
KW Ribosomal protein; Zinc-finger; Metal-binding; Complete proteome.
FT ZN_FING 20 CA-TYPE
SQ SEQUENCE 65 AA; 7176 MW; 56C7E39AC8670AC0 CRC64;
Query Match 60.7%; Score 34; DB 1; Length 65;
Best Local Similarity 70.0%; Pred. No. 5.4;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
OY 1 ADIPEPTGAK 10
DB 44 ATLVEPTGK 53
RESULT 8
RS27_PYPHO STANDARD; PRT; 65 AA.
ID RS27_PYPHO
AC P58078;
DT 16-OCT-2001 (Rel. 40, Created)
DR 16-OCT-2001 (Rel. 40, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE 30S ribosomal protein S27e.
GN RPS27E OR PH1939.1.
OS Pyrococcus horikoshii.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
CC Pyrococcus.
NCBI_TaxID=53953;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-OT3;
RX MEDLINE-98344137; PubMed-9679194;
RA Kawarabayashi Y., Sawada M., Horikawa H., Hino Y.,
RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohnuki Y.,
RA Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,
RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
RA Masuchi Y., Shizuya H., Kikuchi H.;
RT "Complete sequence and gene organization of the genome of a hyper-
thermophilic archaeobacterium, Pyrococcus horikoshii OT3."
RL DNA Res. 5:55-76(1998).
RN [2]
RP IDENTIFICATION.
RA Medigue C., Bocs S.;
RL Unpublished observations (MAY-2001).
CC -1- MISCELLANEOUS: BINDS ONE ZINC ION PER MOLECULE (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE S27E FAMILY OF RIBOSOMAL PROTEINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
DR EMBL: AF000007; -. NOT_ANNOTATED_CDS.
DR InterPro: IPR000592; Ribosomal_S27E.
DR Pfam: PF01667; Ribosomal_S27e; 1.
DR ProDom: PD004466; Ribosomal_S27E; 1.
DR PROSITE: PS01168; RIBOSOMAL_S27E; 1.
KW Ribosomal protein; Zinc-finger; Metal-binding; Complete proteome.
FT ZN_FING 20 CA-TYPE
SQ SEQUENCE 65 AA; 7149 MW; 518078C0A8670AD2 CRC64;
Query Match 60.7%; Score 34; DB 1; Length 65;

Best Local Similarity 70.0%; Pred. No. 5.4;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 ADVEPTGAK 10
1 11111 1
DB 44 ARLVETGSK 53

RESULT 9
ZPRL_SCHPO

ID ZPRL_SCHPO STANDARD; PRT; 459 AA.
AC 013724: 074193; 38. Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE zinc-finger protein zp1.
GN ZPRL OR SPAC15A10.04C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomyces.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069476; PubMed=9852145;
RA Gangwani L., Mikrut M., Galcheva-Gargova Z., Davis R.J.;
RT "Interaction of ZPRL with translation elongation factor-1alpha in
RT proliferating cells.";
RL J. Cell Biol. 143:1471-1484(1998).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Felwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinovitch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volkart G., Aert R., Robben J., Gymnopoulos S.,
RA Wellens I., Vastrelles E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langner I., Beck A., Leinrich H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Medler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Lucibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gallardin C., Tallada V.A., Garçon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerutti L., Lowe T., McCombie M.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).
-1- SUBCELLULAR LOCATION: CYTOPLASMIC; TRANSLocate TO THE NUCLEUS
-1- AFTER NUTRIENT STIMULATION (BY SIMILARITY).
-1- SIMILARITY: BELONGS TO THE ZPRL FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib.ch).

CC EMBL: AF019768: AAC3515.1;
CC EMBL: Z97208: CAB10101.1; -
CC DR

DR InterPro; IPR004457; Znf_ZPRL.
DR Pfam; PF03367; ZPRL; 2.
DR TrRfam; TRFR00310; ZPRL_znf; 2.
KW Nuclear protein; zinc-finger.

FT ZN_FING 38 70 CA-TYPE.
FT ZN_FING 259 291 S->T (IN REF. 1).
FT CONFLICT 88 88 S->A (IN REF. 1).
FT CONFLICT 367 367 G->A (IN REF. 1).
SQ SEQUENCE 459 AA; 51376 MW; ED68F7F02B86F93A CRC64;

Query Match 60.7%; Score 34; DB 1; Length 459;
Best Local Similarity 50.0%; Pred. No. 44;

Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 1 ADVEPTGAK 10
1 1 1 1 1 1
DB 80 ARLVETGSK 89

RESULT 10

ID SOX6_HUMAN STANDARD; PRT; 828 AA.
AC P35712; Q9BX05; Q9BX04; Q9BX03; Q9H018;
DT 01-JUN-1994 (Rel. 29, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Transcription factor SOX-6.

GN SOX6.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND TISSUE SPECIFICITY.

RC TISSUE=lymphocytes, and myoblasts;
RX MEDLINE=21153434; PubMed=1125018;
RA Cohen-Barak O., Hagivara N., Arlt M.F., Horton J.P., Brilliant M.H.;
RT "Cloning, characterization and chromosome mapping of the human SOX6
RT gene.";
RL Gene 265:157-164(2001).
[2]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC TISSUE=Testis;
RX MEDLINE=21154917; PubMed=11230166;
RA Wiemann S., Weil B., Wellenreuther R., Gassenhuber J., Glassl S.,
RA Ansgor W., Boecker M., Blocker H., Bauersachs S., Blum H.,
RA Lauber J., Duesthoeft A., Beyer A., Koehrer K., Strack N.,
RA Mewes H.-W., Oltersweiler B., Obermaier B., Tampe J., Heubner D.,
RA Wambutt R., Korn B., Klein M., Poustka A.;
RT "Towards a catalog of human genes and proteins: sequencing and
RT analysis of 500 novel complete protein coding human cDNAs.";
RL Genome Res. 11:422-435(2001).

RP SEQUENCE OF 632-685 FROM N.A.
RX MEDLINE=92310993; PubMed=1614875;
RA Denby P., Swift S., Brand N., Dabnede N., Barton P., Ashworth A.;
RT "A conserved family of genes related to the testis determining gene,
RT SRY.";
RL Nucleic Acids Res. 20:2887-2887(1992).
-1- FUNCTION: BINDS SPECIFICALLY TO THE DNA SEQUENCE 5'-AACAAAT-3'.
-1- SUBCELLULAR LOCATION: Nuclear (Potential).
-1- ALTERNATIVE PRODUCTS: 3 isoforms; 1 (shown here), 2, and 3; are
-1- produced by alternative splicing.
-1- TISSUE SPECIFICITY: Expressed in a wide variety of tissues, most
-1- abundantly in skeletal muscle.
-1- SIMILARITY: CONTAINS 1 HMG BOX.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC

```
CC or sendan email to license@sb-sib.ch).
CC -----
DR EMBL AF309034 AAK6115.1 -
DR EMBL AF309476 AAK26243.1 JOINED.
DR EMBL AF309471 AAK26243.1 JOINED.
DR EMBL AF309472 AAK26243.1 JOINED.
DR EMBL AF309473 AAK26243.1 JOINED.
DR EMBL AF309474 AAK26243.1 JOINED.
DR EMBL AF309475 AAK26243.1 JOINED.
DR EMBL AF309476 AAK26244.1 -
DR EMBL AF309471 AAK26244.1 JOINED.
DR EMBL AF309472 AAK26244.1 JOINED.
DR EMBL AF309473 AAK26244.1 JOINED.
DR EMBL AF309474 AAK26244.1 JOINED.
DR EMBL AF309475 AAK26244.1 JOINED.
DR EMBL AF309476 CAB66714.1 ALT_INIT.
DR EMBL X65663 CAA46614.1 -
DR PIR S21481 S21481.
DR PIR S22940 S22940.
DR PIR S21485 S21486.
DR HSSP Q05066 IHRT.
DR Genew: HGNC:16421; SOX6.
DR InterPro: IPR000910; HMG_12_box.
DR Pfam: PF00505; HMG_box; 1.
DR SMART: SMO0398; HMG; 1.
KW DNA-binding; Nuclear protein; Transcription regulation; Coiled coil;
KW Alternative splicing.
KW COILED COIL (POTENTIAL).
FT FT DNA_BIND 184 262
FT FT DOMAIN 621 689
FT FT DOMAIN 240 243
FT FT DOMAIN 280 285
FT FT DOMAIN 313 317
FT FT DOMAIN 514 518
FT FT VASPLIC 327 367
FT FT VASPLIC 579 598
FT FT CONFLICT 477 477
FT FT CONFLICT 633 633
SO SEQUENCE 828 AA; 91893 MW; 56CAVCODEA81IDSD CRC64;
Query Match 60.7%; Score 34; DB 1; Length 828;
Best Local Similarity 70.0%; Pred. No. 82;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0.
QY 1 ADVEPTGAK 10
Db 395 AGTVSPTGIK 404
RESULT 11
VP2_ROTBR STANDARD: PRT; 880 AA.
AC PI2472; Q86225;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE RNA-binding protein VP2 (Major internal core protein).
GN S2.
OS Bovine rotavirus (strain RF).
OC Viruses; dsRNA viruses; Reoviridae; Rotavirus.
OX NCBI_TaxID=10933;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89183621; PubMed=2538805;
RA Kumar A., Charpilienne A., Cohen J.;
RT "Nucleotide sequence of the gene encoding for the RNA binding protein
   (VP2) of RF bovine rotavirus."
RL Nucleic Acids Res. 17:2126-2126(1989).
CC -I- FUNCTION: IS THE MOST ABUNDANT CORE PROTEIN. FORMS ICOSAHEDRAL
   STRUCTURES AND BINDS NUCLEIC ACIDS.
CC -I- SUBCELLULAR LOCATION: NUCLEOCAPSID.
CC -I- SIMILARITY: BELONGS TO THE ROTAVIRUS VP2 PROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright It is produced through a collaboration
```

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC of the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/sib.ch).
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; X14057; CAA32213.1; -
DR EMBL; X14057; CAA32215.1; ALT_SEQ.
DR PIR; S03601; S03601.
KW RNA-binding; Nucleocapsid.
FT DOMAIN 536 557 LEUCINE-ZIPPER 1.
FT DOMAIN 665 686 LEUCINE-ZIPPER 2.
SQ SEQUENCE 880 AA; 102496 MW; F4AFABECA7291412 CRC64;

Query Match 60.7%; Score 34; DB 1; Length 880;
Best Local Similarity 77.8%; Pred. No. 88;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ADVEPTGA 9
||||| 1
Db 860 ADVEPTGA 868

RESULT 12
VP2_ROTBU STANDARD: PRT; 881 AA.
ID VP2_ROTBU
AC P17462;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE RNA-binding protein VP2 (Major internal core protein).
DE S2.
OS Bovine rotavirus (strain UK).
OC Viruses; dsRNA viruses; Reoviridae; Rotavirus.
CX NCBI_TaxID=10934;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90326551; PubMed=2165258;
RA Tian Y., Tarlow O., McCrae M.A.;
RT "Nucleotide sequence of gene 2 of the UK tissue culture adapted
RL Nucleic Acids Res. 18:4015-4015(1990).
CC -1- FUNCTION: IS THE MOST ABUNDANT CORE PROTEIN. FORMS ICOSAEDRAL
CC STRUCTURES AND BINDS NUCLEIC ACIDS.
CC -1- SUBCELLULAR LOCATION: NUCLEOCAPSID.
CC -1- SIMILARITY: BELONGS TO THE ROTAVIRUS VP2 PROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC of the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/sib.ch).
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; X52589; CAA36825.1; -
DR PIR; S10252; P2XRUK.
KW RNA-binding; Nucleocapsid.
FT DOMAIN 537 558 LEUCINE-ZIPPER 1 (BY SIMILARITY).
FT DOMAIN 665 688 LEUCINE-ZIPPER 2 (BY SIMILARITY).
SQ SEQUENCE 881 AA; 102487 MW; 7CFDFF653E4CF2CC CRC64;

Query Match 60.7%; Score 34; DB 1; Length 881;
Best Local Similarity 77.8%; Pred. No. 88;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ADVEPTGA 9
||||| 1
Db 861 ADVEPTGA 869

RESULT 13

```

```
VP2_ROTSL
ID VP2_ROTSL STANDARD: PRT: 881 AA.
AC P2672;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE RNA-binding protein VP2 (Major internal core protein).
GN S2.
OS Simian 11 rotavirus (strain SA11-Both).
OC Viruses; dsRNA viruses; Reoviridae; Rotavirus.
OX NCBI_TaxID=37137;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90281597; PubMed=2162107;
RA Mitchell D.B., Both G.W.;
RT "Completion of the genomic sequence of the simian rotavirus SA11:
nucleotide sequences of segments 1, 2, and 3."
RL Virology 177:324-331(1990).
CC -1- FUNCTION: IS THE MOST ABUNDANT CORE PROTEIN. FORMS ICOSAEDRAL
STRUCTURES AND BINDS NUCLEIC ACIDS.
CC -1- SIMILARITY: BELONGS TO THE ROTAVIRUS VP2 PROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
```

```
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
```

Job time : 8.09677 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 18, 2002, 15:53:18 ; Search time 25.1935 Seconds
(without alignments)
89.964 Million cell updates/sec

Title: US-09-848-781-3
Perfect score: 56
Sequence: 1 ADVPEPTGAKE 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPREMBL_21.*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	46	82.1	479	17	Q9UMR2
2	40	71.4	468	17	O57853
3	40	71.4	479	17	O9V2G3
4	39	69.6	847	16	O8Y0C4
5	38	67.9	178	2	O9AGV6
6	38	67.9	307	16	O8XK10
7	37	66.1	211	16	O8Z004
8	37	66.1	211	16	O8Z714
9	37	66.1	219	16	O8YV14
10	37	66.1	236	16	O8X8G3
11	37	66.1	374	4	O9H7X3
12	37	66.1	2417	5	O9U5Z2
13	37	66.1	2417	5	O9XW10
14	36	64.3	159	17	O9HPV8
15	36	64.3	252	16	O8ZC10
16	36	64.3	369	5	O9V1V6

17	36	64.3	515	5	O96002	O96002 drosophila
18	36	64.3	534	16	O808H9	O808H9 agrobacteri
19	36	64.3	612	5	O17894	O17894 caenorhabd1
20	36	64.3	642	16	O59997	O59997 synechocyst
21	36	64.3	1892	5	O97007	O97007 leishmania
22	36	64.3	4834	4	O95714	O95714 homo sapien
23	35	62.5	156	9	O9ZM25	O9ZM25 mycobacteri
24	35	62.5	260	2	O8VPE2	O8VPE2 pseudomonas
25	35	62.5	260	16	O916R1	O916R1 pseudomonas
26	35	62.5	304	16	O9KUC7	O9KUC7 vibrio chol
27	35	62.5	310	2	O47735	O47735 enterococcu
28	35	62.5	318	4	O754Z9	O754Z9 homo sapien
29	35	62.5	335	16	O96EP4	O96EP4 rhizobium 1
30	35	62.5	586	10	O94HX6	O94HX6 oryza sativ
31	35	62.5	686	5	O9V550	O9V550 drosophila
32	35	62.5	902	13	O8UMC5	O8UMC5 gallus gall
33	35	62.5	959	16	O9RL12	O9RL12 streptomyce
34	35	62.5	1025	10	O9SAAS	O9SAAS arabidopsis
35	35	62.5	1224	5	O93330	O93330 caenorhabd1
36	35	62.5	1323	16	O8Y1G6	O8Y1G6 ralstonia s
37	35	62.5	1484	5	O94Z98	O94Z98 caenorhabd1
38	35	62.5	1541	3	O9HE51	O9HE51 neurospora
39	34	60.7	63	17	O8U474	O8U474 pyrococcus
40	34	60.7	120	4	O96HT6	O96HT6 homo sapien
41	34	60.7	153	2	O8VR74	O8VR74 escherichia
42	34	60.7	204	16	O8Z1O9	O8Z1O9 yersinia pe
43	34	60.7	223	16	P74052	P74052 synechocyst
44	34	60.7	254	5	O21223	O21223 caenorhabd1
45	34	60.7	326	5	O9NM65	O9NM65 leishmania

ALIGNMENTS

RESULT 1

Q9UMR2 ID Q9UMR2 PRELIMINARY; PRT; 479 AA.
AC Q9UMR2;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Replication factor C large subunit.
GN RECL OR PF0092.
OS Pyrococcus furiosus.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxID=2261;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VCI;
RX MEDLINE=21172879; PubMed=11274122;
RA Cann I K O., Ishino S., Yuasa M., Daiyasu H., Toh H., Ishino Y.;
RT "Biochemical Analysis of Replication Factor C from the
RT Hyperthermophilic Archaeon Pyrococcus furiosus.";
RL J. Bacteriol. 183:2614-2623(2001).
RL [2]
RP SEQUENCE FROM N.A.
RC STRAIN=VCI / DSM 3638 / ATCC 43587 / JCM 8422;
RA Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.;
RT "The complete sequence of the Pyrococcus furiosus genome.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB034755; BAA8155.1; -;
DR EMBL; AE010135; AAL80216.1; -;
DR InterPro: IPR003593; AAA_ATPase.
DR InterPro: IPR003593; AAA_ATPase-centr.
DR InterPro: IPR000862; RFCDomain.
DR Pfam: PF00004; AAA; 1.
DR SMART; SM00382; AAA; 1.
KW ATP-binding; Complete proteome.
SO SEQUENCE 479 AA; 55287 MW; D3D2092F2D9230E8 CRC64;
Query Match 82.1%; Score 46; DB 17; Length 479;
Best Local Similarity 72.7%; Pred. No. 1.2;

Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 ADVERTISEPAGE 11
 ||:|||||
 Db 110 ADNIEPSGAP 120

RESULT 2
 057853
 ID 057853 PRELIMINARY; PRT; 468 AA.
 AC 057853;
 DT 01-AUG-1998 (TREMblrel. 07, Created)
 DT 01-AUG-1998 (TREMblrel. 07, Last sequence update)
 DE 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE 468AA long hypothetical replication factor C subunit.
 GN PH0113.
 OS Pyrococcus horikoshii.
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
 PY Pyrococcus.
 NCBI_TaxID=53953;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-OT3;
 RX MEDLINE=98344137; PubMed=9679194;
 RA Kawarabayashi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,
 RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
 RA Sekai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,
 RA Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kushiida N., Oguchi A.,
 RA Aoki K.-I., Shizuya H., Nakamura Y., Robb F.T., Horikoshi K.,
 RA Masuchi Y., Shizuya H., Kikuchi H.,
 RT "Complete sequence and gene organization of the genome of a hyper-
 thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";
 RL EMBL: AP000001; BAA29182.1; -;
 DR InterPro: IPR003593; AAA_ATPase.
 DR InterPro: IPR003593; AAA_ATPase_cent.
 DR InterPro: IPR000862; RFCDomain.
 DR Pfam: PF00004; AAA; 1.
 DR SMART: SM00382; AAA; 1.
 KW Complete proteome.
 SQ SEQUENCE 468 AA; 53913 MW; B85739AD2760B5AF CRC64;

Query Match
 Best Local Similarity 71.4%; Score 40; DB 17; Length 468;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 ADVERTISEPAGE 11
 ||:|||||
 Db 110 ADNIEPSGAP 120

RESULT 3
 09V2G3
 ID 09V2G3 PRELIMINARY; PRT; 479 AA.
 AC 09V2G3;
 DT 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
 DE 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE Activator 1, replication factor C, large S UNBONT.
 GN RECL OR PAB0069.
 OS Pyrococcus abyssi.
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
 PY Pyrococcus.
 NCBI_TaxID=29292;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ORSAY;
 RA Heilig R.,
 RT "Pyrococcus abyssi genome sequence: insights into archaeal chromosome
 structure and evolution.";
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ248283; CAB49035.1; -;
 DR InterPro: IPR003593; AAA_ATPase.

DR InterPro: IPR003593; AAA_ATPase_cent.
 DR InterPro: IPR000862; RFCDomain.
 DR Pfam: PF00004; AAA; 1.
 DR SMART: SM00382; AAA; 1.
 KW Complete proteome.
 SQ SEQUENCE 479 AA; 53300 MW; A29BB46E376FB62 CRC64;

Query Match
 Best Local Similarity 71.4%; Score 40; DB 17; Length 479;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 ADVERTISEPAGE 11
 ||:|||||
 Db 110 ADNIEPSGAP 120

RESULT 4
 08Y0C4
 ID 08Y0C4 PRELIMINARY; PRT; 847 AA.
 AC 08Y0C4;
 DT 01-MAR-2002 (TREMblrel. 20, Created)
 DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)
 DE Putative transporter DNA uptake transmembrane protein.
 DE RSC1120 OR RS04634.
 OS Ralstonia solanacearum (Pseudomonas solanacearum).
 OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
 OC Ralstonia.
 NCBI_TaxID=305;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-GM1100;
 RX MEDLINE=21681879; PubMed=11823852;
 RA Selaoudat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
 RA Ariat M., Billault A., Brottier P., Camus J.C., Catolico L.,
 RA Chandler M., Choisme N., Claudel-Renard C., Cunnac S., Demange N.,
 RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schlex T.,
 RA Signier P., Thebaud P., Whalen M., Wincker P., Levy M.,
 RA Weissenbach J., Boucher C.A.,
 RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";
 RL Nature 415:497-502(2002).
 DR EMBL: AL646062; CAD14822.1; -;
 DR InterPro: IPR004477; COMEC_N-term.
 DR InterPro: IPR004797; COMEC_N-term.
 DR TIGRPFAM: TIGR00360; COMEC_N-term; 1.
 DR TIGRPFAM: TIGR00361; COMEC_N-term; 1.
 KW Complete proteome.
 SQ SEQUENCE 847 AA; 90736 MW; 5E664B6CE8685694 CRC64;

Query Match
 Best Local Similarity 69.6%; Score 39; DB 16; Length 847;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 ADVERTISEPAGE 10
 ||:|||||
 Db 837 AETTEPTGAQ 846

RESULT 5
 09AGV6
 ID 09AGV6 PRELIMINARY; PRT; 178 AA.
 AC 09AGV6;
 DT 01-JUN-2001 (TREMblrel. 17, Created)
 DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
 DE Hypothetical 19.4 kDa protein.
 OS Corynebacterium ammoniagenes (Brevibacterium ammoniagenes).
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacteriaceae; Corynebacteriidae;
 OC Corynebacterium.
 NCBI_TaxID=1697;
 RN [1]
 RP SEQUENCE FROM N.A.

```

RC STRAIN=ATCC6872;
RA Lee B., Lee S.H., Suh J., Kim J.H.;
RT "Cloning and DNA sequence of the gene encoding the major sigma factor
   from Corynebacterium ammoniagenes.";
RL Submitted (DEC-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF326348; AAK15631.1; -.
DR HSSP: P20586; 1107.
DR InterPro: IPR000733; Flav_monooxygenase.
DR InterPro: IPR003042; Rng_moxxygenase.
DR Pfam: PF01360; Monooxygenase: 1.
DR PRINTS: PR00420; RNCMNNOXGNASE.
KW Hypothetical protein.
SQ
SEQUENCE 178 AA; 19427 MW; 006FE35A513E80F5 CRC64;

Query Match          67.9%; Score 38; DB 2; Length 178;
Best Local Similarity 80.0%; Pred. No. 15;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ADTVEPTGAK 10
   1 1 1 1 1 1 1
Db 150 AHVTPPTGAK 159

RESULT 6
08XKLO PRELIMINARY; PRT; 307 AA.
AC 08XKLO;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Hypothetical protein CPE1384.
GN CPE1384.
OS Clostridium perfringens.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC Clostridiales; Clostridiaceae; Clostridium.
OX NCBI_TaxID=1502;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=13 / TYPE A;
RX PubMed-11792842;
RA Shimizu T., Ohtani K., Hirakawa H., Ohshima K., Yamashita A.,
   Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.;
RT "Complete genome sequence of Clostridium perfringens, an anaerobic
   flesh-eater.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).
DR EMBL: AP003190; BAB81090.1; -.
KW Hypothetical protein; Complete proteome.
SQ
SEQUENCE 307 AA; 35328 MW; 0E266A01438A2EAF CRC64;

Query Match          67.9%; Score 38; DB 16; Length 307;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 DIVEPTG 8
   1 1 1 1 1 1 1
Db 190 DIVEPTG 196

RESULT 7
08Z004 PRELIMINARY; PRT; 211 AA.
AC 08Z004;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Putative transcriptional repressor (Tetr/AcrR family).
GN tcrG OR STM1213.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed-11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
   Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
   Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Milvaney E.,
   Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
   Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
   LT2.";
RL Nature 413:852-856(2001).
DR EMBL: AE008752; AAL20142.1; -.
DR InterPro: IPR001647; HTH_Tetr.
DR Pfam: PF00440; tetr. 1.
DR PRINTS: PR00455; HTHTEPR.
KW Hypothetical protein; Complete proteome.
SQ
SEQUENCE 211 AA; 23693 MW; DE945085828989EC CRC64;

Query Match          66.1%; Score 37; DB 16; Length 211;
Best Local Similarity 80.0%; Pred. No. 29;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ADTVEPTGAK 10
   1 1 1 1 1 1 1
Db 43 ADLVEATGAK 52

RESULT 8
08Z714 PRELIMINARY; PRT; 211 AA.
AC 08Z714;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Putative Tetr-family regulatory protein.
GN STY1253.
OS Salmonella typhi.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=601;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CT18;
RX MEDLINE=21534947; PubMed-11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Main J.,
   Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
   Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
   Rancin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
   Felwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagers K.,
   Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
   Quail M., Rutherford K., Simmonds M., Skellon J., Stevens K.,
   Whitehead S., Barrrell B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
   enterica serovar Typhi CT18.";
RL Nature 413:848-852(2001).
DR EMBL: AL627269; CAD08337.1; -.
DR InterPro: IPR001647; HTH_Tetr.
DR Pfam: PF00440; tetr. 1.
DR PRINTS: PR00455; HTHTEPR.
KW Hypothetical protein; Complete proteome.
SQ
SEQUENCE 211 AA; 23675 MW; C480B70578CDB94C CRC64;

Query Match          66.1%; Score 37; DB 16; Length 211;
Best Local Similarity 80.0%; Pred. No. 29;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ADTVEPTGAK 10
   1 1 1 1 1 1 1
Db 43 ADLVEATGAK 52

RESULT 9
08YVLA PRELIMINARY; PRT; 219 AA.
ID 08YVLA

```

```

AC Q8YVL4;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Hypothetical protein A1r1961.
GN A1R1961.
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostoc.
OX NCB1_TaxID=103690;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21595285; PubMed=11759940;
RA Kaneo T., Nakamura Y., Wolk C.P., Sasamoto S.,
RA Watanabe A., Iriuch M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
RT cyanobacterium Anabaena sp. strain PCC 7120.";
RL DNA Res. 8:205-213(2001).
DR EMBL: AP003587; BAB73660.1;
DR InterPro: IPR004143; BPL_LiPA_LiPB.
DR Pfam: PF03099; BPL_LiPA_LiPB.1.
DR Hypothetical protein; Complete proteome.
SQ SEQUENCE 219 AA; 24374 MW; F629B/CBAAAAC19D CRC64;

Query Match
Best Local Similarity 80.0%; Score 37; DB 16; Length 219;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ADTVEPTGAK 10
Db 128 ADLVEPTGAK 137
II IIIII
RESULT 10
O8X8G3 PRELIMINARY; PRT; 236 AA.
AC O8X8G3;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE YcfQ protein.
GN YCFQ OR Z1750 OR ECS1489.
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCB1_TaxID=83334;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=O157:H7 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G., III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Postel G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamotis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
DR EMBL: AE005320; AAG5857.1;
DR EMBL: AP002555; BAB34912.1;

```

```

DR InterPro: IPR001647; HTH_Tetr.
DR Pfam: PF00440; tetr. 1.
DR PRINTS: PR00455; HTHETR.
KM Complete proteome.
SQ SEQUENCE 236 AA; 26178 MW; 6D59C4EB1A4A45F CRC64;

Query Match
Best Local Similarity 80.0%; Score 37; DB 16; Length 236;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ADTVEPTGAK 10
Db 69 ADLVEATGAK 78
II IIIII
RESULT 11
O9H7X3 PRELIMINARY; PRT; 374 AA.
AC O9H7X3;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE CDNA FLJ114129 fis, clone MAMMA1002617, weakly similar to zinc finger
DE protein 135.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCB1_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=MAMMARY GLAND;
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshikawa Y.,
RA Matsunawa H., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A.,
RA Nakamura Y., Nagahari K., Masuno Y., Sasaki N.;
RT "NEO human cDNA sequencing project."
RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL: AK024191; BAB14850.1;
DR HSSP: P07248; IARD.
DR InterPro: IPR000822; Znf_C2H2.
DR Pfam: PF00096; Zf-C2H2; 9.
DR PRINTS: PR00048; ZNCFINGER.
DR PRODOM: PD000003; Znf_C2H2; 6.
DR SMART: SM00355; ZNF_C2H2; 9.
DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; 8.
DR PROSITE: PS0157; ZINC_FINGER_C2H2_2; 9.
DR PROSITE: PS0157; ZINC_FINGER_C2H2_2; 9.
KW DNA-binding; Nuclear protein; Zinc-finger.
SQ SEQUENCE 374 AA; 40522 MW; 68B4D485A3FC7EF CRC64;

Query Match
Best Local Similarity 100.0%; Score 37; DB 4; Length 374;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 EPTGAK 11
Db 6 EPTGAK 12
II IIIII
RESULT 12
O9U522 PRELIMINARY; PRT; 2417 AA.
AC O9U522;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Hypothetical 266.8 kDa protein.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Pelodietinae; Caenorhabditis.
OX NCB1_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.

```


RX MEDLINE=20408884; PubMed=10950924;
 RA Guipponi M., Brunschwig K., Chamoun Z., Scott H.S., Shibuya K.,
 RA Kudoh J., Delezoide A.L., El Samadi S., Chetoui Z., Rossier C.,
 RA Shimizu N., Mueller F., Delabar J.M., Antonakis S.E.;
 RT "C21orf1, a novel human chromosome 21 gene, has a caenorhabditis
 RT elegans ortholog (pad-1) required for embryonic patterning.";
 RL Genomics 68:30-40(2000).
 DR EMBL: AJ250261; CAB5919.1; -;
 KW Hypothetical protein.
 SQ SEQUENCE 2417 AA; 266850 MW; 2238DDE2E0E0F390 CRC64;

Query Match 66.1%; Score 37; DB 5; Length 2417;
 Best Local Similarity 87.5%; Pred. No. 4.3e+02;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 DVEPTGA 9
 | | | | |
 DB 1077 DNEPTGA 1084

RESULT 13

O9XW10 PRELIMINARY; PRT; 2417 AA.

AC O9XW10; 01-NOV-1999 (TREMblrel. 12, Created)
 DT 01-OCT-2001 (TREMblrel. 18, last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, last annotation update)
 DE Y18D10A.13 protein.
 GN Y18D10A.13.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 RX NCBL_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RL Harris B.R.;
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99069613; PubMed=9851916;
 RA none;
 RT "Genome sequence of the nematode C.elegans: A platform for
 RT investigating biology.";
 RT Science 282:2012-2018(1998).
 DR EMBL: AL034393; CAA22324.2; -;
 SQ SEQUENCE 2417 AA; 267150 MW; DAC7D4F9A5249E06 CRC64;

Query Match 66.1%; Score 37; DB 5; Length 2417;
 Best Local Similarity 87.5%; Pred. No. 4.3e+02;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 DVEPTGA 9
 | | | | |
 DB 1077 DNEPTGA 1084

RESULT 14

O9HPV8 PRELIMINARY; PRT; 159 AA.

AC O9HPV8; 01-MAR-2001 (TREMblrel. 16, Created)
 DT 01-MAR-2001 (TREMblrel. 16, last sequence update)
 DT 01-MAR-2001 (TREMblrel. 16, last annotation update)
 DE Vng1453H.
 GN VNG1453H.
 OS Halobacterium sp. (strain NRC-1).
 OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
 OC Halobacteriaceae; Halobacterium.
 OX NCBL_TaxID=64091;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20504483; PubMed=11016950;
 RA Ng W.V., Kennedy S.P., Mahairas G.G., Bergquist B., Pan M.,

RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Shrogha J.,
 RA Swartzell S., Weir D., Hall J., Dahl T.A., Welter R., Goo Y.A.,
 RA Leitzauser B., Keller K., Cruz R., Danson M.J., Houh D.W.,
 RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
 RA Isehnarger T.A., Peck R.F., Pohlschroder M., Spudich J.L., Jung K.-H.,
 RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
 RA Eshardt H., Lowe T.W., Liang P., Riley M., Hood U., Dassarma S.;
 RT "Genome sequence of Halobacterium species NRC-1.";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
 DR EMBL: AE005061; AAC19759.1; -;
 KW Complete proteome.
 SQ SEQUENCE 159 AA; 17080 MW; 3BA5079FDCBA157 CRC64;

Query Match 64.3%; Score 36; DB 17; Length 159;
 Best Local Similarity 60.0%; Pred. No. 33;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 ADVTEPTGAK 10
 | | | | |
 DB 141 ADSLEPTGPR 150

RESULT 15

O8ZC10 PRELIMINARY; PRT; 252 AA.

ID O8ZC10; 01-MAR-2002 (TREMblrel. 20, Created)
 DT 01-MAR-2002 (TREMblrel. 20, last sequence update)
 DT 01-MAR-2002 (TREMblrel. 20, last annotation update)
 DE Hypothetical protein YPO3220.
 GN YPO3220.
 OS Versinia pestis.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Versinia.
 OX NCBL_TaxID=632;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CO-92 / BIOVAR ORIENTALIS;
 RX MEDLINE=21470413; PubMed=11586360;
 RA Parkhill J., Wren B.W., Thomson N.R., Tittball R.W., Holden M.T.G.,
 RA Prentice M.B., Sebahia M., James K.D., Churcher C., Mungall K.L.,
 RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Farraga A.M.,
 RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan A.V.,
 RA Felwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
 RA Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,
 RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;
 RT "Genome sequence of Versinia pestis, the causative agent of plague.";
 RL Nature 413:523-527(2001).
 DR EMBL: A014156; CAC92455.1; -;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 252 AA; 28526 MW; 0BFDF0FD63C0494E CRC64;

Query Match 64.3%; Score 36; DB 16; Length 252;
 Best Local Similarity 100.0%; Pred. No. 55;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 VEPTGAK 10
 | | | | |
 DB 81 VEPTGAK 87

Search completed: November 18, 2002, 15:58:12
 Job time : 27.1935 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 18, 2002, 15:53:18 ; Search time 31.9355 Seconds
(without alignments)
45.897 Million cell updates/sec

Title: US-09-848-781-2
Perfect score: 64
Sequence: 1 ACDEYLQYKE 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A.Geneseq.101002:*

- 1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
- 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
- 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
- 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
- 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
- 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
- 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
- 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
- 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
- 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
- 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
- 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
- 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
- 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
- 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
- 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
- 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
- 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
- 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
- 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
- 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
- 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
- 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	64	100.0	11	23	AAU78380
2	54	84.4	289	23	ABB47676
3	40	62.5	18	22	ABB29133
4	40	62.5	18	22	ABB34293
5	40	62.5	18	22	AAU03050
6	40	62.5	18	22	ABG37094
7	39	60.9	1438	22	ABG03631
8	39	60.9	1438	22	ABG28308
9	39	60.9	1439	22	ABG14849
10	38	59.4	219	22	ABB71413

11	37	57.8	70	22	ABG23412
12	37	57.8	89	21	AAG26667
13	37	57.8	89	21	AAG37739
14	37	57.8	92	21	AAG26666
15	37	57.8	92	21	AAG37738
16	37	57.8	95	21	AAG26665
17	37	57.8	95	21	AAG37737
18	37	57.8	508	22	ABG17426
19	36	56.2	26	22	ABG35043
20	36	56.2	26	22	ABG20453
21	36	56.2	26	22	AAU68228
22	36	56.2	26	23	ABG37760
23	36	56.2	115	19	AAU86000
24	36	56.2	121	23	ABP25605
25	36	56.2	137	22	AAU42999
26	36	56.2	382	22	ABG19988
27	36	56.2	538	22	AAU67476
28	36	56.2	2050	20	AAU90117
29	36	56.2	2050	20	AAU73499
30	36	56.2	2050	21	AAU67241
31	36	56.2	2813	7	AAU60462
32	36	56.2	2813	19	AAU54347
33	36	56.2	2813	21	AAU70557
34	36	56.2	2813	23	AAU75317
35	35.5	55.5	689	22	ABB63086
36	35	54.7	22	21	AAU87217
37	35	54.7	22	22	AAE06199
38	35	54.7	22	23	ABG34023
39	35	54.7	69	14	AAU39128
40	35	54.7	89	22	AAU90685
41	35	54.7	104	23	AAU06168
42	35	54.7	132	22	AAU00628
43	35	54.7	139	23	ABP03334
44	35	54.7	147	23	ABG53327
45	35	54.7	186	23	ABB90205

ALIGNMENTS

RESULT 1	AAU78380	
ID	AAU78380	standard; Peptide; 11 AA.
XX	AAU78380;	
AC	18-JUN-2002	(first entry)
XX		
DT		
XX		
DE		Bacterial detection method associated peptide PI.
XX		
KW		Bacterial detection method; pathogen; food product; food retail;
KW		metalloprotease; gastroenteritis; food-related work area;
KW		fluorescence resonance energy transfer; FRET.
XX		
OS		Synthetic.
XX		
FH	Key	Location/Qualifiers
FT	Modified-site	1
FT		/label= DABSYL
FT	Modified-site	11
FT		/note= "Fluorescent group"
FT		/label= EDANS
FT		/note= "Fluorescent group"
PN	MO200210433-A2.	
XX		
PD	07-FEB-2002.	
XX		
PF	03-MAY-2001; 2001MO-US14613.	
XX		
PR	03-MAY-2000; 2000US-201405P.	
XX		
PA	(EXPR-) EXPRESSIVE CONSTRUCTS INC.	

Novel human diagno
Arabidopsis thalia
Arabidopsis thalia
Arabidopsis thalia
Arabidopsis thalia
Arabidopsis thalia
Arabidopsis thalia
Arabidopsis thalia
Novel human diagno
Peptide #2549 enco
Protein #2452 enco
Human bone marrow
Human peptide enco
S. pneumoniae deri
Streptococcus poly
Propionibacterium
Drosophila melanog
Amino acid sequenc
Human mature von W
Human von Willebra
Human von Willebra
Sequence of human
Canine von Willebr
Canine von Willebr
Human von Willebra
Drosophila melanog
Human secreted pro
Human gene 65 enco
Human secreted pro
LD78 Phe12>Asp. H
C glutamitum prote
Human ORF protein
Human polypeptide
Human ORF protein
Lactococcus lactis
Human polypeptide

XX Sanders MC;
PI
XX
XX WPI; 2002-303940/34.
DR
XX
PT Detection of bacterial contamination in foods or food-related work
PT areas comprises identifying a protein specific to the microorganism
PT using fluorescence or colorimetric methods -
PS
XX Disclosure; Page 7; 25pp; English.
XX
CC The invention describes a specific and sensitive method for detection of
CC a pathogenic microorganism in potentially contaminated food products
CC at the retail level, by detecting a bacteria-specific protein e.g. a
CC listeria monocytogenes metalloprotease. The process is useful for
CC detecting listeria monocytogenes, a common cause of gastroenteritis, or
CC other food contaminants in food products or food-related work areas.
CC This sequence represents the peptide PI that functions as a substrate
CC for a broad range of bacterial pathogens. The peptide can be used as a
CC probe for fluorescence resonance energy transfer (FRET) analysis of
CC bacterial protease activity in food products.
XX
SQ Sequence 11 AA;
Query Match 100.0%; Score 64; DB 23; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.00088;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 ACCDEYLQTFKE 11
DB 1 ACCDEYLQTFKE 11
RESULT 2
ABBA7676
ID ABBA7676 standard; protein; 289 AA.
XX
AC ABBA7676;
XX
DT 05-FEB-2002 (first entry)
XX
DE Listeria monocytogenes protein #380.
XX
KW Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
KW vitamin B12; bacterial infection; disease.
XX
OS Listeria monocytogenes.
XX
PN WO200177335-A2.
XX
PD 18-OCT-2001.
XX
PF 11-APR-2001; 2001WO-FR01118.
XX
PR 11-APR-2000; 2000FR-0004629.
XX
PA (INSP) INST PASTEUR.
XX
PI Buchrieser C, Frangoul L, Couve E, Rusnick C, Fajbi H, Dehoux P;
PI Dussurget O, Chetoui F, Nedjari H, Glaser P, Kunst F, Cossart P;
PI Daniels J, Goebel W, Kreft J, Kuhn M, Ng E, Vazquez-Boland JA;
PI Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A;
PI Chakraborty T, Dommann E, Hain T, Berche P, Charbit A, Durant L;
PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N;
PI Madueno E, De Pablos B, Wehland J, Kaerst U, Entian K, Haut J;
PI Rose M, Voss H;
XX
DR WPI; 2002-010914/01.
XX
PT Genomic sequence for listeria monocytogenes, useful e.g. for treatment
PT and prevention of listeria and related bacterial infections, and
PT related polypeptides -
XX

PS Claim 6; SEQ ID No 381; 192pp; French.
XX
XX The present invention relates to the genome sequence of listeria
CC monocytogenes EGD-e (see AB03041). The genome sequence and fragments of
CC it are useful for selecting probes and primers for detecting genes in L.
CC monocytogenes and related organisms, and for studying genetic
CC polymorphisms and other genomes. The present sequence is a protein
CC encoded by the genome sequence of the present invention. Proteins
CC expressed from the genome sequence are useful for raising specific
CC antibodies. Identification of L. monocytogenes and related organisms
CC for biosynthesis and biodegradation, especially biosynthesis of vitamin
CC B12. The genome sequence and proteins encoded by it are also useful for
CC selecting compounds that regulate gene expression and cell replication
CC and modulate L. monocytogenes-related diseases. In addition, the genome
CC sequence and proteins encoded by it are useful in pharmaceutical and
CC vaccines compositions for the treatment or prevention of infections by L.
CC monocytogenes and related organisms.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pot_sequences.
XX
SQ Sequence 289 AA;
Query Match 84.4%; Score 54; DB 23; Length 289;
Best Local Similarity 100.0%; Pred. No. 0.7;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 ACCDEYLQTF 9
DB 27 ACCDEYLQTF 35
RESULT 3
ABB29133
ID ABB29133 standard; peptide; 18 AA.
XX
AC ABB29133;
XX
DT 01-FEB-2002 (first entry)
XX
DE Peptide #1784 encoded by breast cell single exon nucleic acid probe.
XX
KW Human; microarray; single exon probe; gene expression; breast;
KW disease; cancer.
XX
OS Homo sapiens.
XX
PN WO200157271-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00662.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-496933/54.
XX
XX New spatially-addressable set of single exon nucleic acid probes,
PT useful for measuring gene expression in sample derived from human
PT breast, comprises number of single exon nucleic acid probes -
PS
XX Claim 27; SEQ ID NO 12101; 327pp + sequence listing; English.
XX

CC The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human breast and BT 474 cells. The method involves contacting
CC the probes with a collection of detectably labelled nucleic acids
CC derived from mRNA of human breast, and then measuring the label
CC bound to each probe of the microarray. The probes are useful for
CC verifying the expression of regions of genomic DNA predicted to
CC encode proteins. They are useful for gene discovery, and for
CC determining predisposition and/or prognosing breast disease. Gene
CC expression analysis is useful for assessing the toxicity of chemical
CC agents on cells. The microarray of this invention presents a far greater
CC diversity of probes for measuring gene expression, with far less bias
CC than expressed sequence tag microarrays. The method is suitable for
CC rapid production of functional information from genomic sequence. The
CC present sequence is a peptide encoded by a single exon nucleic acid
CC probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 18 AA;

Query Match 62.5%; Score 40; DB 22; Length 18;

Best Local Similarity 63.6%; Pred. No. 8.6;

Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 ACCDEYLQTKR 11
||||: |||

DB 4 ACCDDLTLTAE 14

RESULT 4
ABB34293
ID ABB34293 standard; Peptide; 18 AA.

AC ABB34293;

DT 04-FEB-2002 (first entry)

DE Peptide #1799 encoded by human foetal liver single exon probe.

KW Human; foetal liver; gene expression; single exon nucleic acid probe.

OS Homo sapiens.

PN WO200157277-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00669.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-483447/52.

CC Human genome-derived single exon nucleic acid probes useful for
CC analyzing gene expression in human fetal liver -
CC Claim 27; SEQ ID NO 26928; 639pp + sequence listing; English.
CC The invention relates to a single exon nucleic acid probe for
CC measuring human gene expression in a sample derived from human foetal
CC liver. The single exon nucleic acid probes may be used for predicting,

CC measuring and displaying gene expression in samples derived from human
CC fetal liver. The present sequence is a peptide encoded by a single exon
CC nucleic acid probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 18 AA;

Query Match 62.5%; Score 40; DB 22; Length 18;

Best Local Similarity 63.6%; Pred. No. 8.6;

Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 ACCDEYLQTKR 11
||||: |||

DB 4 ACCDDLTLTAE 14

RESULT 5
AAM03050
ID AAM03050 standard; Protein; 18 AA.

AC AAM03050;

DT 09-OCT-2001 (first entry)

DE Peptide #1732 encoded by probe for measuring breast gene expression.

KW Probe; human; breast disease; breast cancer; development disorder;

KW Inflammatory disease; proliferative breast disease; non-carcinoma tumour.

OS Homo sapiens.

PN WO200157270-A2.

PD 09-AUG-2001.

PF 29-JAN-2001; 2001WO-US00661.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-476286/51.

CC Novel single exon nucleic acid probe used to measuring gene expression
CC in a human breast -
CC Claim 27; SEQ ID NO 11790; 322pp; English.

CC The present invention relates to novel single exon nucleic acid probes
CC (see A110010-A1110067). The present sequence is a peptide encoded by one
CC such probe. The probes are useful for measuring human gene expression in
CC a human breast sample, where the probe hybridises at high stringency to a
CC nucleic acid expressed in the human breast. The probes are useful for
CC predicting, diagnosing, grading, staging, monitoring and prognosing
CC diseases of the human breast, particularly those diseases with polygenic
CC aetiology. The diseases include: breast cancer, disorders of development,
CC inflammatory diseases of the breast, fibrocystic changes, proliferative
CC breast disease and non-carcinoma tumours.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 18 AA;

Query Match 62.5%; Score 40; DB 22; Length 18;
Best Local Similarity 63.6%; Pred. No. 8.6;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 ACCDEYLQTK 11
Db 4 ACCDDLTLTAE 14

RESULT 6
ABG37094
ID ABG37094 standard; Peptide; 18 AA.
XX
AC ABG37094;
XX
DT 19-AUG-2002 (first entry)
XX
DE Human peptide encoded by genome-derived single exon probe SEQ ID 26759.
XX
KW Human; single exon probe; asthma; lung cancer; COPD; ILD;
KW chronic obstructive pulmonary disease; interstitial lung disease;
KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW primary ciliary dyskinesia; pulmonary hypertension;
KW hyaline membrane disease.
XX
OS Homo sapiens.
XX
PN WO200186003-A2.
XX
PD 15-NOV-2001.
XX
PE 30-JAN-2001; 2001MO-US00665.
XX
PR 04-FEB-2000; 2000US-180312P.
XX
PR 26-MAY-2000; 2000US-207456P.
XX
PR 30-JUN-2000; 2000US-0608408.
XX
PR 03-AUG-2000; 2000US-0632366.
XX
PR 21-SEP-2000; 2000US-234687P.
XX
PR 27-SEP-2000; 2000US-236359P.
XX
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2002-114183/15.
XX
XX
PT Spatially-addressable set of single exon nucleic acid probes, used to
measure gene expression in human lung samples -
XX
XX
PS Claim 27; SEQ ID NO 26759; 634bp; English.
XX
XX
CC The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human lung comprising single exon nucleic acid probes having one of
CC 12614 nucleic acid sequences mentioned in the specification, or their
CC complements or the 12387 open reading frames derived from the 12614
CC probes. Also included are a microarray comprising the novel set of
CC probes; the novel set of probes which hybridize at high stringency to a
CC nucleic acid expressed in the human lung; measuring gene expression to a
CC sample derived from human lung, comprising (a) contacting the array with
CC a collection of detectably labeled nucleic acids derived from human lung
CC mRNA, and (b) measuring the label detectably bound to each probe of
CC the array; identifying exons in a eukaryotic genome, comprising
CC (a) algorithmically predicting at least one exon from genomic sequences
CC of the eukaryote; and (b) detecting specific hybridisation of detectably
CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC having a fragment identical to the predicted exon, the probe is included

CC In the above mentioned microarray, assigning exons to a single gene,
CC comprising (a) identifying exons from genomic sequence by the method
CC above and (b) measuring the expression of each of the exons in several
CC tissues and/or cell types using hybridisation to a single exon
CC microarrays having a probe with the exon, where a common pattern of
CC expression of the exons in the tissues and/or cell types indicates that
CC the exons should be assigned to a single gene; a peptide comprising one
CC of 12011 sequences, mentioned in the specification, or encoded by the
CC probes/open reading frames (ORF). The probes are used for gene
CC expression analysis, and for identifying exons in a gene, particularly
CC using human lung derived mRNA and for the study of lung diseases
CC such as asthma, lung cancer, chronic obstructive pulmonary disease
CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
CC haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
CC pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic
CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
CC and hyaline membrane disease. The present sequence is a peptide/protein
CC encoded by a single exon probe of the invention.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC http://wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 18 AA;
XX

Query Match 62.5%; Score 40; DB 23; Length 18;
Best Local Similarity 63.6%; Pred. No. 8.6;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 ACCDEYLQTK 11
Db 4 ACCDDLTLTAE 14

RESULT 7
ABG03631
ID ABG03631 standard; Protein; 1438 AA.
XX
AC ABG03631;
XX
DT 13-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #3622.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PE 30-MAR-2001; 2001MO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
XX
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
XX
DR N-PSDB; AAS67818.
XX
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 20; SEQ ID NO 33990; 103bp; English.

CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

Sequence 1439 AA;

Query Match 60.9%; Score 39; DB 22; Length 1439;
 Best Local Similarity 60.0%; Pred. No. 7.3e+02;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 CCDEYLTQKE 11
 1 111 111:
 Db 643 CIDECYQTRD 652

RESULT 10
 ABB71413
 ID ABB71413 standard; Protein; 219 AA.

AC ABB71413;

XX 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 41031.

XX Drosophila: developmental biology; cell signalling; insecticide;

KM pharmaceutical.

XX Drosophila melanogaster.

OS WO200171042-A2.

PN 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

PE 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PMD, Myers EW;

DR WPI: 2001-656860/75.

XX N-PSDB; ABL15516.

PT New isolated nucleic acid detection reagent for detecting 1000 or more
 CC genes from Drosophila and for elucidating cell signalling and cell-cell
 CC interactions -

PS Disclosure: SEQ ID NO 41031; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA

CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (ABB57737-ABB72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

Sequence 219 AA;

Query Match 59.4%; Score 38; DB 22; Length 219;
 Best Local Similarity 55.6%; Pred. No. 1.8e+02;
 Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 1 ACCDEYLTQ 9
 111111:
 Db 211 ACCEDHLOS 219

RESULT 11
 ABG23412
 ID ABG23412 standard; Protein; 70 AA.

XX ABG23412;

DT 18-FEB-2002 (first entry)

DE Novel human diagnostic protein #23403.

XX Human: chromosome mapping; gene mapping; gene therapy; forensic;
 KM food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US08631.

PE 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YF;

DR WPI: 2001-639362/73.

XX N-PSDB; AAS87599.

PT New isolated polynucleotide and encoded polypeptides, useful in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits and to assess
 CC biodiversity -

PS Claim 20; SEQ ID NO 53771; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.

CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 70 AA;

Query Match 57.8%; Score 37; DB 22; Length 70;

Best Local Similarity 62.5%; Pred. No. 91;

Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 2 CCDEYLOT 9

Db 33 CCDDWLST 40

RESULT 12

AC AAG26667 standard; Protein: 89 AA.

XX AAG26667;

DT 17-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 31209.

KW Protein identification; signal transduction pathway; metabolic pathway;

KW hybridisation assay; genetic mapping; gene expression control; promoter;

XX termination sequence.

OS Arabidopsis thaliana.

PN EP1033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 28-APR-1999; 99US-0130891.

PR 30-APR-1999; 99US-0131449.

PR 04-MAY-1999; 99US-0132407.

PR 05-MAY-1999; 99US-0132484.

PR 06-MAY-1999; 99US-0132485.

PR 07-MAY-1999; 99US-0132487.

PR 11-MAY-1999; 99US-0132863.

PR 18-JUN-1999; 99US-0139454.

PR 18-JUN-1999; 99US-0139455.

PR 18-JUN-1999; 99US-0139456.

PR 18-JUN-1999; 99US-0139457.

PR 18-JUN-1999; 99US-0139458.

PR 18-JUN-1999; 99US-0139459.

PR 18-JUN-1999; 99US-0139460.

PR 18-JUN-1999; 99US-0139461.

PR 18-JUN-1999; 99US-0139462.

PR 18-JUN-1999; 99US-0139463.

PR 18-JUN-1999; 99US-0139750.

PR 21-JUN-1999; 99US-0139817.

PR 22-JUN-1999; 99US-0139899.

PR 23-JUN-1999; 99US-0140353.

PR 24-JUN-1999; 99US-0140354.

PR 28-JUN-1999; 99US-0140695.

PR 29-JUN-1999; 99US-0140823.

PR 30-JUN-1999; 99US-0140991.

PR 01-JUL-1999; 99US-0141287.

PR 01-JUL-1999; 99US-0141842.

PR 02-JUL-1999; 99US-0142154.

PR 06-JUL-1999; 99US-0142390.

PR 08-JUL-1999; 99US-0142803.

PR 09-JUL-1999; 99US-0142820.

PR 12-JUL-1999; 99US-0142977.

PR 13-JUL-1999; 99US-0143542.

PR 14-JUL-1999; 99US-0143624.

PR 15-JUL-1999; 99US-0144005.

PR 16-JUL-1999; 99US-0144085.

PR 16-JUL-1999; 99US-0144086.

PR 19-JUL-1999; 99US-0144325.

PR 19-JUL-1999; 99US-0144331.

PR 19-JUL-1999; 99US-0144332.

PR 19-JUL-1999; 99US-0144333.

PR 19-JUL-1999; 99US-0144334.

PR 19-JUL-1999; 99US-0144335.

PR 20-JUL-1999; 99US-0144352.

PR 20-JUL-1999; 99US-0144632.

PR 20-JUL-1999; 99US-0144634.

PR 21-JUL-1999; 99US-0144814.

PR 21-JUL-1999; 99US-0145086.

PR 21-JUL-1999; 99US-0145088.

PR 22-JUL-1999; 99US-0145089.

PR 22-JUL-1999; 99US-0145087.

PR 22-JUL-1999; 99US-0145089.

PR 23-JUL-1999; 99US-0145192.

PR 23-JUL-1999; 99US-0145145.

PR 23-JUL-1999; 99US-0145218.

PR 23-JUL-1999; 99US-0145224.

PR 26-JUL-1999; 99US-0145276.

PR 27-JUL-1999; 99US-0145913.

PR 27-JUL-1999; 99US-0145918.

PR 28-JUL-1999; 99US-0145919.

PR 28-JUL-1999; 99US-0145951.

PR 02-AUG-1999; 99US-0146386.

PR 02-AUG-1999; 99US-0146388.

PR 03-AUG-1999; 99US-0146389.

PR 04-AUG-1999; 99US-0147038.

PR 04-AUG-1999; 99US-0147204.

PR 05-AUG-1999; 99US-0147302.

PR 05-AUG-1999; 99US-0147192.

PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156586.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 23-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match

57.88; Score 37; DB 21; Length 89;

Best Local Similarity 50.08; Pred. No. 1.1e+02;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
Oy 2 CCDEYLOTFKE 11
|||::||
Db 46 CCDDKLEOFKE 55
RESULT 13
AAG37739
ID AAG37739 standard; Protein; 89 AA.
XX AC
XX AAG37739;
XX DT
XX 18-OCT-2000 (first entry)
XX DE
XX Arabidopsis thaliana protein fragment SEQ ID NO: 46453.
XX KW
XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX OS
XX Arabidopsis thaliana.
XX PN
XX EP1033405-A2.
XX PD
XX 06-SEP-2000.
XX PF
XX 25-FEB-2000; 2000EP-0301439.
XX XX
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 23-APR-1999; 99US-0130891.
XX 28-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 30-APR-1999; 99US-0132407.
XX 04-MAY-1999; 99US-0132484.
XX 05-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.
XX 06-MAY-1999; 99US-0132487.
XX 07-MAY-1999; 99US-0132863.
XX 11-MAY-1999; 99US-0134256.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134219.
XX 14-MAY-1999; 99US-0134221.
XX 14-MAY-1999; 99US-0134370.
XX 18-MAY-1999; 99US-0134768.
XX 19-MAY-1999; 99US-0134941.
XX 20-MAY-1999; 99US-0135124.
XX 21-MAY-1999; 99US-0135353.
XX 24-MAY-1999; 99US-0135629.
XX 25-MAY-1999; 99US-0136021.
XX 27-MAY-1999; 99US-0136392.
XX 28-MAY-1999; 99US-0136782.
XX 01-JUN-1999; 99US-0137222.
XX 03-JUN-1999; 99US-0137528.
XX 04-JUN-1999; 99US-0137502.
XX 07-JUN-1999; 99US-0137724.
XX 08-JUN-1999; 99US-0138094.
XX 10-JUN-1999; 99US-0138540.
XX 10-JUN-1999; 99US-0138847.
XX 14-JUN-1999; 99US-0139119.

PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142053.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 10-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.

PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0158029.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 22-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 23-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 57 8%; Score 37; DB 21; Length 89;
Best Local Similarity 50.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 2 CDEYLQTRK 11
||: : : ||
DB 46 CCDKLIEQKE 55

RESULT 14
AAG26666
ID AAG26666 standard; Protein; 92 AA.
XX
AC AAG26666;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 31208.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence.
XX
OS Arabidopsis thaliana.
PN EP1033405-A2.
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-01332407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0135421.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.

PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142820.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144332.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0144816.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.

PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 57.8%; Score 37; DB 21; Length 92;

Best Local Similarity 50.0%; Pred. No. 1.2e+02; Mismatches 2; Indels 0; Gaps 0;

Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 2 CCDEYLQTKK 11
Db 49 CCCKLIEQKE 58

RESULT 15
AAG37738
ID AAG37738 standard; Protein; 92 AA.
XX
AC AAG37738;
XX

DT 18-OCT-2000 (first entry)
XX Arabidopsis thaliana protein fragment SEQ ID NO: 46452.
DE
XX Protein identification: signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX Arabidopsis thaliana.
OS
XX
PN EP1033405-A2.
PD
XX 06-SEP-2000.
PF
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134321.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.

PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142290.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145813.
PR 27-JUL-1999; 99US-0145818.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.

PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-015138.
PR 01-SEP-1999; 99US-015130.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 57.8%; Score 37; DB 21; Length 92;
Best Local Similarity 50.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
QY 2 CCDEYLQTRK 11
Db 49 CCDDKLEQKE 58

Search completed: November 18, 2002, 15:55:26
Job time : 33.9355 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compen Ltd.

OM protein - protein search, using sw model

Run on: November 18, 2002, 15:53:18 ; Search time 12.7742 Seconds
(without alignments)
25.336 Million cell updates/sec

Title: US-09-848-781-2
Perfect score: 64
Sequence: 1 ACDEYIQTKE 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 segs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCITUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/Backfilest1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	36	56.2	36	1	US-08-451-472-8
2	36	56.2	2050	2	US-08-347-594A-2
3	36	56.2	2813	3	US-08-896-449A-2
4	36	56.2	2813	3	US-09-132-652-2
5	35	54.7	644	2	US-08-866-757-2
6	35	54.7	644	4	US-09-153-593-2
7	35	54.7	1024	4	US-09-562-737-81
8	35	54.7	1024	4	US-09-562-737-82
9	34	53.1	182	4	US-08-858-207A-434
10	34	53.1	305	4	US-09-315-794-22
11	34	53.1	305	4	US-09-389-341-22
12	34	53.1	766	5	PCR-US94-00198-6
13	34	53.1	1024	4	US-09-562-737-86
14	34	53.1	1024	4	US-09-562-737-90
15	34	53.1	1719	2	US-08-459-568-4
16	34	53.1	1719	3	US-08-399-411-4
17	34	53.1	1719	3	US-08-516-859A-4
18	34	53.1	1719	4	US-09-586-472-4
19	34	53.1	1719	4	US-09-528-706-4
20	33	51.6	255	3	US-09-009-913-9
21	33	51.6	265	3	US-09-009-913-7
22	33	51.6	446	4	US-09-134-001C-5125
23	33	51.6	1024	4	US-09-562-737-89
24	32	50.0	300	4	US-09-293-549-14
25	32	50.0	395	2	US-08-841-349-9
26	32	50.0	583	1	US-08-448-196A-5
27	32	50.0	615	2	US-08-663-566A-17

28	32	50.0	615	2	US-08-023-610-17	Sequence 17, Appl
29	32	50.0	615	2	US-08-288-065A-17	Sequence 17, Appl
30	32	50.0	615	2	US-08-362-240A-17	Sequence 17, Appl
31	32	50.0	615	5	PCR-US95-10245-17	Sequence 17, Appl
32	32	50.0	642	4	US-09-337-307A-3	Sequence 3, Appl
33	32	50.0	642	4	US-09-337-307A-4	Sequence 4, Appl
34	32	50.0	1162	2	US-08-663-566A-15	Sequence 15, Appl
35	32	50.0	1162	2	US-08-023-610-15	Sequence 15, Appl
36	32	50.0	1162	2	US-08-288-065A-15	Sequence 15, Appl
37	32	50.0	1162	2	US-08-362-240A-15	Sequence 15, Appl
38	32	50.0	1162	5	PCR-US95-10245-15	Sequence 15, Appl
39	32	50.0	1872	1	US-08-188-582-14	Sequence 14, Appl
40	32	50.0	1872	1	US-08-646-715-14	Sequence 14, Appl
41	32	50.0	1893	1	US-08-188-582-11	Sequence 11, Appl
42	32	50.0	1893	1	US-08-646-715-11	Sequence 11, Appl
43	31	49.2	499	4	US-09-336-643A-8	Sequence 8, Appl
44	31	48.4	102	1	US-08-340-203A-12	Sequence 12, Appl
45	31	48.4	102	2	US-08-452-427-12	Sequence 12, Appl

ALIGNMENTS

RESULT 1
US-08-451-472-8
; Sequence 8, Application US/08451472
; Patent No. 5770192
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: BIOLOGICAL CONTROL AGENTS
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon & Vanderhye PC
; STREET: 8th Floor, 1100 No. 5770192th Glebe Road
; CITY: Arlington
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/451,472
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/117,125
; FILING DATE: 24-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Mary J Wilson
; REGISTRATION NUMBER: 32,955
; REFERENCE/DOCKET NUMBER: 117-187
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816 4000
; TELEFAX: (703) 816 4100
; TELEEX: 200797 NIXN UR
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 36 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-451-472-8

Query Match 56.2%; Score 36; DB 1; Length 36;
Best local Similarity 100.0%; Pred. No. 15;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CCDEY 6
DB 20 CCDEY 24

RESULT 2
US-08-347-594A-2
Sequence 2, Application US/08347594A
Patent No. 5849536
GENERAL INFORMATION:
APPLICANT: Garfinkel, Leonard
TITLE OF INVENTION: CLONING AND PRODUCTION OF HUMAN VON
WILLEBRAND FACTOR GPTD BINDING DOMAIN POLYPEPTIDES AND
METHODS OF USING SAME
TITLE OF INVENTION: WILLEBRAND FACTOR GPTD BINDING DOMAIN POLYPEPTIDES AND
METHODS OF USING SAME
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESSES:
ADDRESSEE: John P. White
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/347,594A
FILING DATE: No. 5849536ember 30, 1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 36537-B2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-278-0400
TELEFAX: 212-391-0525
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2050 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-347-594A-2
Query Match
Best Local Similarity 56.2%; Score 36; DB 2; Length 2050;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 CCDEY 6
Db 1625 CCDEY 1629
RESULT 3
US-08-896-449A-2
Sequence 2, Application US/08896449A
Patent No. 6040143
GENERAL INFORMATION:
APPLICANT: Venta, Patrick J
APPLICANT: Yuzbaslyan-Gurkan, Vilma
APPLICANT: Schall, William D
APPLICANT: Brewer, George J
TITLE OF INVENTION: DNA ENCODING CANINE VON WILLEBRAND
FACTOR AND METHODS OF USE
TITLE OF INVENTION: FACTOR AND METHODS OF USE
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Harness, Dickey & Pierce, P.L.C.
STREET: 5445 Corporate Drive
CITY: Troy
STATE: Michigan
COUNTRY: USA
ZIP: 48098

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/896,449A
FILING DATE: 18-JUL-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Smith, Deann F.
REFERENCE/DOCKET NUMBER: 2115-001226
TELECOMMUNICATION INFORMATION:
TELEPHONE: 248-641-1600
TELEFAX: 248-641-0270
TELEX: 287637
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2813 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-896-449A-2
Query Match
Best Local Similarity 56.2%; Score 36; DB 3; Length 2813;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 CCDEY 6
Db 2388 CCDEY 2392
RESULT 4
US-09-132-652-2
Sequence 2, Application US/09132652
Patent No. 6074832
GENERAL INFORMATION:
APPLICANT: Venta, Patrick J
APPLICANT: Yuzbaslyan-Gurkan, Vilma
APPLICANT: Schall, William D
APPLICANT: Brewer, George J
APPLICANT: Duffendeck, John
TITLE OF INVENTION: DNA ENCODING CANINE VON WILLEBRAND FACTOR AND METHODS
OF USE
FILE REFERENCE: 2115S-001226CPB
CURRENT APPLICATION NUMBER: US/09/132,652
EARLIER APPLICATION NUMBER: 08/896,449
EARLIER FILING DATE: 1997-07-18
NUMBER OF SEQ ID NOS: 29
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 2
LENGTH: 2813
TYPE: PRT
ORGANISM: Canis familiaris
US-09-132-652-2
Query Match
Best Local Similarity 56.2%; Score 36; DB 3; Length 2813;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 CCDEY 6
Db 2388 CCDEY 2392
RESULT 5
US-08-866-757-2
Sequence 2, Application US/08866757
Patent No. 5858716
GENERAL INFORMATION:
APPLICANT: ELISHOURBAGY, NABIL A

APPLICANT: LI, XIAOTONG
APPLICANT: BERGSM, DERK J
TITLE OF INVENTION: NOVEL 7TM RECEPTOR (H2CAAT1)
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: RATNER & PRESTIA
STREET: P.O. BOX 980
CITY: VALLEY FORGE
STATE: PA
COUNTRY: USA
ZIP: 19482
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/866,757
FILING DATE: 30-MAY-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: PRESTIA, PAUL F
REGISTRATION NUMBER: 23,031
REFERENCE/DOCKET NUMBER: GH-70055
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-407-0700
TELEFAX: 610-407-0701
TELEX: 846169
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 644 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-866-757-2

Query Match 54.7%; Score 35; DB 2; Length 644;
Best Local Similarity 55.6%; Pred. No. 3.5e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 CCDEYLOTK 10
||:| | |
DB 553 CCESFLTK 561

RESULT 6
US-09-153-593-2
; Sequence 2, Application US/09153593A
; Patent No. 6174994
; GENERAL INFORMATION:
; APPLICANT: ELSHOURBAGY, NABIL A
; APPLICANT: LI, XIAOTONG
; APPLICANT: BERGSM, DERK J
; TITLE OF INVENTION: NOVEL 7TM RECEPTOR (H2CAAT1)
; FILE REFERENCE: GH-70055-1
; CURRENT APPLICATION NUMBER: US/09/153,593A
; CURRENT FILING DATE: 1998-09-15
; EARLIER APPLICATION NUMBER: 08/866,757
; EARLIER FILING DATE: 1997-05-30
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 644
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-09-153-593-2

Query Match 54.7%; Score 35; DB 4; Length 644;
Best Local Similarity 55.6%; Pred. No. 3.5e+02;

Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 CCDEYLOTK 10
||:| | |
DB 553 CCESFLTK 561

RESULT 7
US-09-562-737-81
; Sequence 81, Application US/09562737
; Patent No. 6428967
; GENERAL INFORMATION:
; APPLICANT: Herz, Joachim
; APPLICANT: Gotthardt, Michael
; TITLE OF INVENTION: LDL Receptor Signaling Pathways
; FILE REFERENCE: UTSW0708
; CURRENT APPLICATION NUMBER: US/09/562,737
; CURRENT FILING DATE: 2000-05-01
; NUMBER OF SEQ ID NOS: 132
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 81
; LENGTH: 1024
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-562-737-81

Query Match 54.7%; Score 35; DB 4; Length 1024;
Best Local Similarity 55.6%; Pred. No. 5.4e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 CCDEYLOTK 10
||:| | |
DB 693 CCDAWLKV 701

RESULT 8
US-09-562-737-82
; Sequence 82, Application US/09562737
; Patent No. 6428967
; GENERAL INFORMATION:
; APPLICANT: Herz, Joachim
; APPLICANT: Gotthardt, Michael
; TITLE OF INVENTION: LDL Receptor Signaling Pathways
; FILE REFERENCE: UTSW0708
; CURRENT APPLICATION NUMBER: US/09/562,737
; CURRENT FILING DATE: 2000-05-01
; NUMBER OF SEQ ID NOS: 132
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 82
; LENGTH: 1024
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-562-737-82

Query Match 54.7%; Score 35; DB 4; Length 1024;
Best Local Similarity 55.6%; Pred. No. 5.4e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 CCDEYLOTK 10
||:| | |
DB 693 CCDAWLKV 701

RESULT 9
US-08-858-207A-434
; Sequence 434, Application US/08858207A
; Patent No. 6348328

GENERAL INFORMATION:
APPLICANT: Black, Michael
APPLICANT: Hodgson, John
APPLICANT: Knowles, David
APPLICANT: Nicholas, Richard
APPLICANT: Stodola, Robert
TITLE OF INVENTION: No. 6348328el Compounds
NUMBER OF SEQUENCES: 552
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/858, 207A
FILING DATE: 09-MAY-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/017670
FILING DATE: 14-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Gimmil, Edward R.
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: P50475
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ. ID NO: 434:
SEQUENCE CHARACTERISTICS:
LENGTH: 182 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: NO. 6348328e
US-08-858-207A-434

Query Match 53.1%; Score 34; DB 4; Length 182;
Best Local Similarity 60.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 CDEYLOTKE 11
DB 57 CCGSLEYE 66

RESULT 10
US-09-315-794-22
Sequence 22, Application US/09315794
Patent No. 6197517
GENERAL INFORMATION:
APPLICANT: Roberts, Christopher J.
TITLE OF INVENTION: ESSENTIAL GENES OF YEAST AS TARGETS FOR ANTIFUNGAL
TITLE OF INVENTION: AGENTS, HERBICIDES, INSECTICIDES AND ANTI-PROLIFERATION
FILE REFERENCE: 9301-053
CURRENT APPLICATION NUMBER: US/09/315,794
CURRENT FILING DATE: 1999-05-21
NUMBER OF SEQ ID NOS: 64
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 22
LENGTH: 305
TYPE: PRT
ORGANISM: Saccharomyces cerevisiae
US-09-315-794-22

Query Match 53.1%; Score 34; DB 4; Length 305;
Best Local Similarity 66.7%; Pred. No. 2.4e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 CDEYLOTKE 11
DB 179 CDELLQNK 187

RESULT 11
US-09-389-341-22
Sequence 22, Application US/09389341
Patent No. 620803
GENERAL INFORMATION:
APPLICANT: Roberts, Christopher J.
TITLE OF INVENTION: ESSENTIAL GENES OF YEAST AS TARGETS FOR ANTIFUNGAL
TITLE OF INVENTION: AGENTS, HERBICIDES, INSECTICIDES AND ANTI-PROLIFERATIVE
FILE REFERENCE: 9301-057
CURRENT APPLICATION NUMBER: US/09/389,341
CURRENT FILING DATE: 1999-09-02
EARLIER APPLICATION NUMBER: 09/315,794
EARLIER FILING DATE: 1999-05-21
NUMBER OF SEQ ID NOS: 72
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 22
LENGTH: 305
TYPE: PRT
ORGANISM: Saccharomyces cerevisiae
US-09-389-341-22

Query Match 53.1%; Score 34; DB 4; Length 305;
Best Local Similarity 66.7%; Pred. No. 2.4e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 CDEYLOTKE 11
DB 179 CDELLQNK 187

RESULT 12
PCT-US94-00198-6
Sequence 6, Application PC/TUS9400198
GENERAL INFORMATION:
APPLICANT: Schering Corp.
TITLE OF INVENTION: RAS Associated GAP Proteins
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Schering Corp.
STREET: 1 Giraldo Farms
CITY: Madison
STATE: New Jersey
COUNTRY: USA
ZIP: 94304-1104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: 6.0.8
SOFTWARE: Microsoft Word 5.1a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/00198
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/004,824
FILING DATE: 15-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Lunn, Paul G.
REGISTRATION NUMBER: 32,743
REFERENCE/DOCKET NUMBER: DX0352 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (201)822-7255
TELEFAX: (201)822-7039

INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 766 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 ORIGINAL SOURCE:
 ORGANISM: Schizosaccharomyces pombe
 PCT-US94-00198-6

Query Match 53.1%; Score 34; DB 5; Length 766;
 Best Local Similarity 66.7%; Pred. No. 5.9e+02;
 Matches 8; Conservative 1; Mismatches 1; Indels 2; Gaps 1;

QY 1 ACCD-EYLOTQ 10
 |||:|:|
 DB 615 AICDHNEYLOTQ 626

RESULT 13
 US-09-562-737-86
 ; Sequence 86, Application US/09562737
 ; Patent No. 6428967
 ; GENERAL INFORMATION:
 ; APPLICANT: Herz, Joachim
 ; APPLICANT: Gotthardt, Michael
 ; TITLE OF INVENTION: LDL Receptor Signaling Pathways
 ; FILE REFERENCE: UTSW0708
 ; CURRENT APPLICATION NUMBER: US/09/562,737
 ; CURRENT FILING DATE: 2000-05-01
 ; NUMBER OF SEQ ID NOS: 132
 ; SOFTWARE: Patentin Ver. 2.1
 ; SEQ ID NO 86
 ; LENGTH: 1024
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
 ; US-09-562-737-86

Query Match 53.1%; Score 34; DB 4; Length 1024;
 Best Local Similarity 55.6%; Pred. No. 7.9e+02;
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 CCDEYLOTK 10
 |||:|:|
 DB 693 CCDGWLKVK 701

RESULT 14
 US-09-562-737-90
 ; Sequence 90, Application US/09562737
 ; Patent No. 6428967
 ; GENERAL INFORMATION:
 ; APPLICANT: Herz, Joachim
 ; APPLICANT: Gotthardt, Michael
 ; TITLE OF INVENTION: LDL Receptor Signaling Pathways
 ; FILE REFERENCE: UTSW0708
 ; CURRENT APPLICATION NUMBER: US/09/562,737
 ; CURRENT FILING DATE: 2000-05-01
 ; NUMBER OF SEQ ID NOS: 132
 ; SOFTWARE: Patentin Ver. 2.1
 ; SEQ ID NO 90
 ; LENGTH: 1024
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
 ; US-09-562-737-90

Query Match 53.1%; Score 34; DB 4; Length 1024;
 Best Local Similarity 55.6%; Pred. No. 7.9e+02;
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 CCDEYLOTK 10
 |||:|:|
 DB 693 CCDGWLKVK 701

RESULT 15
 US-08-459-568-4
 ; Sequence 4, Application US/08459568
 ; Patent No. 5811304
 ; GENERAL INFORMATION:
 ; APPLICANT: Huang, Shi
 ; TITLE OF INVENTION: Retinoblastoma Protein - Interacting
 ; TITLE OF INVENTION: Zinc Finger Proteins
 ; NUMBER OF SEQUENCES: 93
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Campbell and Flores
 ; STREET: 4370 La Jolla Village Drive, Suite 700
 ; CITY: San Diego
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 92122
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/459,568
 ; FILING DATE: 02-JUN-1995
 ; CLASSIFICATION: 514
 ; PRIORITY APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/399,411
 ; FILING DATE: 06-MAR-1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Campbell, Cathryn A.
 ; REGISTRATION NUMBER: 31,815
 ; REFERENCE/DOCKET NUMBER: P-L1 1264
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (619) 535-9001
 ; TELEFAX: (619) 535-8949
 ; INFORMATION FOR SEQ ID NO: 4:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1719 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-459-568-4

Query Match 53.1%; Score 34; DB 2; Length 1719;
 Best Local Similarity 60.0%; Pred. No. 1.3e+03;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 CCDEYLOTKE 11
 |||:|:|
 DB 800 CFDEYKSKKE 809

Search completed: November 18, 2002, 15:56:50
 Job time : 13.7742 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 18, 2002, 15:56:17 ; Search time 6.03226 Seconds
(without alignments)
27.464 Million cell updates/sec

Title: US-09-848-761-2
Perfect score: 64
Sequence: 1 ACCDEYLQTRK 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 97044 seqs, 15060890 residues

Total number of hits satisfying chosen parameters: 97044

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US08_NFM_PUB.pep:*
- 2: /cgn2_6/ptodata/1/pubpaa/PC7_NFM_PUB.pep:*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NFM_PUB.pep:*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NFM_PUB.pep:*
- 6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/1/pubpaa/PC7US_PUBCOMB.pep:*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_NFM_PUB.pep:*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/1/pubpaa/US10_NFM_PUB.pep:*
- 12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
- 13: /cgn2_6/ptodata/1/pubpaa/US60_NFM_PUB.pep:*
- 14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	62.5	18	US-09-864-761-47551	Sequence 47551, A
2	36	56.2	26	US-09-864-761-35751	Sequence 35751, A
3	36	56.2	2813	US-09-381-261A-1	Sequence 1, Appl1
4	36	56.2	2813	US-09-886-900-2	Sequence 2, Appl1
5	35	54.7	260	US-09-925-299-975	Sequence 975, App
6	35	54.7	431	US-09-862-767A-8	Sequence 8, Appl1
7	35	54.7	436	US-09-862-767A-4	Sequence 4, Appl1
8	35	54.7	631	US-09-862-767A-2	Sequence 2, Appl1
9	35	54.7	2045	US-09-736-969A-95	Sequence 95, Appl1
10	35	54.7	2045	US-09-925-301-1186	Sequence 1186, App
11	34	53.1	141	US-09-815-242-13493	Sequence 13493, A
12	34	53.1	252	US-09-815-242-10853	Sequence 10853, A
13	34	53.1	254	US-09-815-242-5018	Sequence 5018, App
14	34	53.1	669	US-09-864-761-38000	Sequence 38000, A
15	34	53.1	983	US-09-864-761-33932	Sequence 33932, A
16	33	51.6	44	US-09-864-761-34400	Sequence 34400, A
17	33	51.6	44	US-09-864-761-48442	Sequence 48442, A
18	33	51.6	190	US-09-815-242-5877	Sequence 5877, App
19	33	51.6	195	US-09-864-761-48442	Sequence 48442, A

20	33	51.6	215	10	US-09-815-242-13061	Sequence 13061, A
21	33	51.6	321	10	US-09-815-242-10934	Sequence 10934, A
22	33	51.6	449	10	US-09-815-242-5375	Sequence 5375, App
23	33	51.6	469	10	US-09-815-242-12290	Sequence 12290, A
24	33	51.6	1382	10	US-09-729-653-2	Sequence 2, Appl1
25	32	50.0	220	10	US-09-864-761-43037	Sequence 43037, A
26	32	50.0	226	10	US-09-815-242-11709	Sequence 11709, A
27	32	50.0	406	10	US-09-731-872-284	Sequence 284, App
28	32	50.0	461	10	US-09-764-864-1288	Sequence 1288, App
29	32	50.0	501	9	US-09-344-882-22	Sequence 22, Appl1
30	32	50.0	636	10	US-09-764-864-832	Sequence 832, App
31	32	50.0	682	10	US-09-815-242-11452	Sequence 11452, A
32	31	48.4	40	10	US-09-250-883-19	Sequence 19, Appl1
33	31	48.4	85	10	US-09-925-297-829	Sequence 829, App
34	31	48.4	103	10	US-09-107-058-3	Sequence 3, Appl1
35	31	48.4	103	10	US-09-761-117-3	Sequence 3, Appl1
36	31	48.4	109	12	US-10-062-254-308	Sequence 308, App
37	31	48.4	121	10	US-09-764-864-1279	Sequence 1279, App
38	31	48.4	126	10	US-09-934-586A-20	Sequence 20, Appl1
39	31	48.4	146	10	US-09-764-864-822	Sequence 822, App
40	31	48.4	176	10	US-09-815-242-13271	Sequence 13271, A
41	31	48.4	245	9	US-09-895-913A-48	Sequence 48, Appl1
42	31	48.4	308	10	US-09-804-682-168	Sequence 168, App
43	31	48.4	340	10	US-09-250-883-17	Sequence 17, Appl1
44	31	48.4	379	10	US-09-804-682-72	Sequence 72, Appl1
45	31	48.4	379	10	US-09-804-682-73	Sequence 73, Appl1

ALIGNMENTS

RESULT 1
US-09-864-761-47551
Sequence 47551, Application US/09864761
Patient No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharron G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEOTIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecm1ca-X-1
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 09/864,761
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30

```

; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 47551
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC011331.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.6
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.2
US-09-864-761-47551
```

```

Query Match          62.5%; Score 40; DB 10; Length 18;
Best Local Similarity 63.6%; Pred. No. 0.53;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 1 ACCDEYLQTKK 11
    |||||
Db 4 ACCDDLLTAE 14
```

```

RESULT 2
US-09-864-761-35751
```

```

; Sequence 35751, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
```

```

; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 35751
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL033532.12
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 4.8
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.99
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.3
US-09-864-761-35751
```

```

Query Match          56.2%; Score 36; DB 10; Length 26;
Best Local Similarity 100.0%; Pred. No. 3.5;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 2 CCDEY 6
    |||||
Db 14 CCDEY 18
```

```

RESULT 3
```

```

US-09-381-261A-1
; Sequence 1, Application US/09381261A
; Patent No. US20020123457A1
; GENERAL INFORMATION:
; APPLICANT: Loscalzo, Joseph
; APPLICANT: Imbal, Aida
; TITLE OF INVENTION: NO. US20020123457A1el Anti-Platelet Agent
; FILE REFERENCE: 102258.327
; CURRENT APPLICATION NUMBER: US/09/381,261A
; CURRENT FILING DATE: 1999-12-07
; PRIOR APPLICATION NUMBER: PCT/US98/06092
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: US 60/046,981
; PRIOR FILING DATE: 1997-03-27
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 2813
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-381-261A-1
```

```

Query Match          56.2%; Score 36; DB 10; Length 2813;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 2 CCDEY 6
    |||||
Db 2388 CCDEY 2392
```

```

RESULT 4
```

```

US-09-886-900-2
; Sequence 2, Application US/09886900
; Patent No. US20020137051A1
; GENERAL INFORMATION:
; APPLICANT: Venta, Patrick J
; APPLICANT: Yuzbasiyan-Gurkan, Vilma
; APPLICANT: Schall, William D
; APPLICANT: Brewer, George J
```

```

; TITLE OF INVENTION: DNA ENCODING CANINE VON WILLEBRAND
; FACTOR AND METHODS OF USE
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Harness, Dickey & Pierce, P.L.C.
; STREET: 5445 Corporate Drive
; CITY: Troy
; STATE: Michigan
; COUNTRY: USA
; ZIP: 48098
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/886,900
; FILING DATE: 21-Jun-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/896,449
; FILING DATE: 18-JUL-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, Deann F.
; REFERENCE/DOCKET NUMBER: 2115-001226
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 248-641-1600
; TELEFAX: 248-641-0270
; TELEX: 287637
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2813 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-886-900-2

Query Match          56.2%; Score 36; DB 10; Length 2813;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CCDEYLQTK 6
DB 2388 CCDEY 2392

RESULT 5
US-09-925-299-975
; Sequence 975, Application US/09925299
; Patent No. US20020055627A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA102
; CURRENT APPLICATION NUMBER: US/09/925,299
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 975
; LENGTH: 260
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (212)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-299-975
```

```

Query Match          54.7%; Score 35; DB 10; Length 260;
Best Local Similarity 55.6%; Pred. No. 45;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 CCDEYLQTK 10
DB 169 CCESFLTK 177

RESULT 6
US-09-862-767A-8
; Sequence 8, Application US/09862767A
; Patent No. US20020034786A1
; GENERAL INFORMATION:
; APPLICANT: Pan, Yang
; TITLE OF INVENTION: TANGO-72 AND TANGO-77 NUCLEIC ACID MOLECULES AND POLYPEPTIDES
; FILE REFERENCE: MB101997-008P1RCPC1N1(M)
; CURRENT APPLICATION NUMBER: US/09/862,767A
; CURRENT FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: US 09/127,856
; PRIOR FILING DATE: 1998-08-03
; PRIOR APPLICATION NUMBER: US 60/054,646
; PRIOR FILING DATE: 1997-08-04
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-862-767A-8

Query Match          54.7%; Score 35; DB 10; Length 431;
Best Local Similarity 55.6%; Pred. No. 72;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 CCDEYLQTK 10
DB 400 CCESFLTK 408

RESULT 7
US-09-862-767A-4
; Sequence 4, Application US/09862767A
; Patent No. US20020034786A1
; GENERAL INFORMATION:
; APPLICANT: Pan, Yang
; TITLE OF INVENTION: TANGO-72 AND TANGO-77 NUCLEIC ACID MOLECULES AND POLYPEPTIDES
; FILE REFERENCE: MB101997-008P1RCPC1N1(M)
; CURRENT APPLICATION NUMBER: US/09/862,767A
; CURRENT FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: US 09/127,856
; PRIOR FILING DATE: 1998-08-03
; PRIOR APPLICATION NUMBER: US 60/054,646
; PRIOR FILING DATE: 1997-08-04
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 436
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-862-767A-4

Query Match          54.7%; Score 35; DB 10; Length 436;
Best Local Similarity 55.6%; Pred. No. 73;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 CCDEYLQTK 10
DB 400 CCESFLTK 408

RESULT 8
US-09-862-767A-2
```

; Sequence 2, Application US/09862767A
; Patent No. US20020034786A1
; GENERAL INFORMATION:
; APPLICANT: Pan, Yang
; TITLE OF INVENTION: TANGO-72 AND TANGO-77 NUCLEIC ACID MOLECULES AND POLYPEPTIDES
; FILE REFERENCE: M810197-008P1C1C1(M)
; CURRENT APPLICATION NUMBER: US/09/862,767A
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: US 09/127,856
; PRIOR FILING DATE: 1998-08-03
; PRIOR APPLICATION NUMBER: US 60/054,646
; PRIOR FILING DATE: 1997-08-04
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 631
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-862-767A-2

Query Match 54.7%; Score 35; DB 10; Length 631;
Best Local Similarity 55.6%; Pred. No. 1e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 CCDEYLOTK 10
11: 1 11
Db 540 CCESFLTK 548

RESULT 9
US-09-736-969A-95
; Sequence 95, Application US/09736969A
; Patent No. US2002006302A1
; GENERAL INFORMATION:
; APPLICANT: Lu, Peter
; APPLICANT: Garman, Jonathan David
; APPLICANT: Candia III, Albert Frederick
; APPLICANT: Arbor Vita Corporation
; TITLE OF INVENTION: CLASP-4 Transmembrane Protein
; FILE REFERENCE: 020054-000411US
; CURRENT APPLICATION NUMBER: US/09/736,969A
; CURRENT FILING DATE: 2000-12-13
; PRIOR APPLICATION NUMBER: US 60/160,860
; PRIOR FILING DATE: 1999-10-21
; PRIOR APPLICATION NUMBER: US 60/162,498
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: US 60/170,453
; PRIOR FILING DATE: 1999-12-13
; PRIOR APPLICATION NUMBER: US 60/176,195
; PRIOR FILING DATE: 2000-01-14
; PRIOR APPLICATION NUMBER: US 60/182,296
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: US 09/547,276
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,267
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,460
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,527
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,528
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 09/687,837
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,503
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,508
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,539
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,543
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 153

; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 95
; LENGTH: 2045
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human CLASP-7
US-09-736-969A-95

Query Match 54.7%; Score 35; DB 10; Length 2045;
Best Local Similarity 50.0%; Pred. No. 3.1e+02;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 2 CCDEYLOTK 11
11: 1 11
Db 354 CCDEYWLKE 363

RESULT 10
US-09-736-960-92
; Sequence 92, Application US/09736960
; Patent No. US20020102267A1
; GENERAL INFORMATION:
; APPLICANT: Lu, Peter
; APPLICANT: Garman, Jonathan David
; APPLICANT: Candia III, Albert Frederick
; APPLICANT: Arbor Vita Corporation
; TITLE OF INVENTION: CLASP-5 Transmembrane Protein
; FILE REFERENCE: 020054-000511US
; CURRENT APPLICATION NUMBER: US/09/736,960
; CURRENT FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: US 60/160,860
; PRIOR FILING DATE: 1999-10-21
; PRIOR APPLICATION NUMBER: US 60/162,498
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: US 60/170,453
; PRIOR FILING DATE: 1999-12-13
; PRIOR APPLICATION NUMBER: US 60/176,195
; PRIOR FILING DATE: 2000-01-14
; PRIOR APPLICATION NUMBER: US 60/182,296
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: US 09/547,276
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,267
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,460
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,527
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,528
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 09/687,837
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,503
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,508
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,539
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,543
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 92
; LENGTH: 2045
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human CLASP-7
US-09-736-960-92

Query Match 54.7%; Score 35; DB 10; Length 2045;
Best Local Similarity 50.0%; Pred. No. 3.1e+02;

Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 2 CCDEYLQTK 11
11:1:11
DB 354 CCEPYVWLKE 363

RESULT 11
US-09-925-301-1186
; Sequence 1186, Application US/09925301
; Patent No. US2002005308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1186
; LENGTH: 141
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-301-1186

Query Match
Best Local Similarity 75.1%; Score 34; DB 10; Length 141;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 3 CDEYLQTK 10
111111
DB 60 CDAYLQMK 67

RESULT 12
US-09-815-242-13493
; Sequence 13493, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13493

; LENGTH: 252
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-815-242-13493

Query Match
Best Local Similarity 53.1%; Score 34; DB 10; Length 252;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 2 CCDEYLQTK 11
111111
DB 127 CCGESLETYE 136

RESULT 13
US-09-815-242-10853
; Sequence 10853, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10853
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-815-242-10853

Query Match
Best Local Similarity 53.1%; Score 34; DB 10; Length 254;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 2 CCDEYLQTK 11
111111
DB 129 CCGETLETYE 138

RESULT 14
US-09-815-242-5018
; Sequence 5018, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.

```

APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELTRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5018
LENGTH: 669
TYPE: PRT
ORGANISM: Enterococcus faecalis
FEATURE:
NAME/KEY: VARIANT
LOCATION: (1)...(669)
OTHER INFORMATION: Xaa = Any Amino Acid
US-09-815-242-5018

```

```

Query Match      53.1%; Score 34; DB 10; Length 669;
Best Local Similarity 60.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```

```

QY      2 CCDEVLYOTKE 11
DB      540 CCGETLETYE 549

```

```

RESULT 15
US-09-864-761-38000
Sequence 38000. Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aemica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30

```

```

PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 38000
LENGTH: 983
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AL031277.1
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.6
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.4
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.6
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.8
OTHER INFORMATION: EST_HUMAN HIT: A0140221.1, EVALUATE 1.00e-111
OTHER INFORMATION: SWISSPROT HIT: Q61116, EVALUATE 1.00e-09
US-09-864-761-38000

```

```

Query Match      53.1%; Score 34; DB 10; Length 983;
Best Local Similarity 60.0%; Pred. No. 2.3e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```

```

QY      2 CCDEVLYOTKE 11
DB      103 CPEYKMSKE 112

```

```

Search completed: November 18, 2002, 16:05:45
Job time : 7.03226 secs

```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 18, 2002, 15:53:18 ; Search time 13.129 Seconds
(without alignments)
80.545 Million cell updates/sec

Title: US-09-848-781-2
Perfect score: 64
Sequence: 1 ACCDEYIQTKE 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 segs, 96134422 residues
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	54	84.4	289	2 S20888	phospholipase C (E
2	54	84.4	289	2 C43868	lecithinase - list
3	54	84.4	289	2 AFI100	phospholipase C 11
4	45	70.3	1111	2 T26972	hypothetical prote
5	41	64.1	560	2 T05083	hypothetical prote
6	40	62.5	767	2 T07693	hypothetical prote
7	39	60.9	227	2 A71676	hypothetical prote
8	38	59.4	252	2 T09635	triose-phosphate i
9	38	59.4	1791	2 T24089	hypothetical prote
10	37	57.8	147	2 S28065	invH protein precu
11	37	57.8	147	2 S28064	invH protein precu
12	37	57.8	147	2 AH0852	cell adherence/inv
13	37	57.8	330	2 D64526	hypothetical prote
14	37	57.8	470	2 S62064	mitochondrial RNA
15	37	57.8	840	2 A42707	type II transposas
16	37	57.8	883	2 A71434	probable RNA helic
17	37	57.8	1235	1 VGIHMU	E2 glycoprotein pr
18	37	57.8	1376	1 VGIHMU	E2 glycoprotein pr
19	37	57.8	1376	1 JQ1534	E2 glycoprotein pr
20	37	57.8	1407	1 BMBYMI	TARA box-binding p
21	37	57.8	4767	2 T31345	hypothetical prote
22	36	56.2	63	1 T1YF	proteinase inhibit
23	36	56.2	96	2 S69007	trypsin/chymotryps
24	36	56.2	114	2 S69006	trypsin/chymotryps
25	36	56.2	123	2 E95240	conserved hypothet
26	36	56.2	123	2 F98104	conserved hypothet
27	36	56.2	441	2 A40550	peroxisomal membra
28	36	56.2	538	1 VGNZMU	cell fusion glycop
29	36	56.2	538	2 S52472	cell fusion protel

30	36	56.2	709	1 P21V50	RNA-directed RNA p
31	36	56.2	714	2 S66699	hypothetical prote
32	36	56.2	813	2 T40622	translation elonga
33	36	56.2	1119	2 A88481	protein C16A3.6 (1
34	36	56.2	1324	1 VGIH59	E2 glycoprotein pr
35	36	56.2	2813	1 VWHU	von Willebrand fac
36	36	56.2	7829	2 T15789	hypothetical prote
37	35.5	55.5	1476	2 A45773	kelch protein, lon
38	35	54.7	243	2 C86628	prophage psi prote
39	35	54.7	243	2 S11226	MyD88 protein - mo
40	35	54.7	247	2 A69952	conserved hypothet
41	35	54.7	302	2 AH1749	transcription regu
42	35	54.7	335	2 T32657	hypothetical prote
43	35	54.7	377	2 C65051	rubredoxin-NAD+ re
44	35	54.7	377	2 AE0845	probable rubredoxi
45	35	54.7	377	2 F85919	probable oxidoredu

ALIGNMENTS

RESULT 1

S20888
phospholipase C (EC 3.1.4.3) precursor - Listeria monocytogenes (fragment)
C:Species: Listeria monocytogenes
C:Date: 16-Sep-1992 #sequence_revision 16-Sep-1992 #text_change 24-Sep-1999
C:Accession: S20888
R:Domann, E.; Weiland, J.; Rohde, M.; Pistor, S.; Hartl, M.; Goebel, W.; Leimweiser-W
EMBO J. 11, 1981-1990, 1992
A>Title: A novel bacterial virulence gene in Listeria monocytogenes required for host
A:Reference number: S20887; MUID:92258410; PMID:1582425
A:Accession: S20888
A:Molecule type: DNA
A:Residues: 1-289 <DOM>
A:Cross-references: EMBL:X59723; NID:g44091; PIDN:CAA42408.1; PID:g44093
C:Genetics:
A:Gene: plcB
C:Superfamily: Bacillus cereus phospholipase C
C:Keywords: phosphoric diester hydrolase
F:1-25/Domain: signal sequence #status predicted <Sig>
F:26-289/Product: phospholipase C #status predicted <Mat>

Query Match 84.4%; Score 54; DB 2; Length 289;
Best Local Similarity 100.0%; Pred. No. 0.076;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCDEYIQT 9
DB 27 ACCDEYIQT 35

RESULT 2

C43868
lecithinase - Listeria monocytogenes
C:Species: Listeria monocytogenes
C:Date: 10-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 24-Sep-1998
C:Accession: C43868
R:Razquez-Boland, J.A.; Kocks, C.; Dramsi, S.; Ohayon, H.; Geoffroy, C.; Mengaud, J.;
Infect. Immun. 60, 219-230, 1992
A>Title: Nucleotide sequence of the lecithinase operon of Listeria monocytogenes and
A:Reference number: A43868; MUID:92104678; PMID:1309513
A:Accession: C43868
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-289 <VAZ>
A>Note: sequence inconsistent with the nucleotide translation
A>Note: sequence extracted from NCBI backbone (NCBIN:74437, NCBI:74462)
C:Superfamily: Bacillus cereus phospholipase C

Query Match 84.4%; Score 54; DB 2; Length 289;
Best Local Similarity 100.0%; Pred. No. 0.076;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ACCDEYLQ7 9
| | | | | | | | | |
Db 27 ACCDEYLQ7 35

RESULT 3

AF1100
phospholipase C [imported] - Listeria monocytogenes (strain EGD-e)
C:Species: Listeria monocytogenes
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C:Accession: AF1100
R:Laaser, P.; Frangoul, L.; Buchtleser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecher, J.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihl, H.; D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kref, J.; Kuhn, M.; Kunst, F.; Kurapk, G.; Madueno, E.; Maltournam, A.; Maok, C.; Schleuter, T.; Simoes, N.; Tjereez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend, A.; Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AF1100
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-289 <GLA>
A:Cross-references: GB:NC_003210; PIDN:CAD00732.1; PID:q16409570; GSPDB:GN00177
A:Experimental source: strain EGD-e
C:Genetics:
A:Gene: PICB
C:Superfamily: Bacillus cereus phospholipase C

Query Match 84.4%; Score 54; DB 2; Length 289;
Best Local Similarity 100.0%; Pred. No. 0.076;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ACCDEYLQ7 9
| | | | | | | | | |
Db 27 ACCDEYLQ7 35

RESULT 4

T26972
hypothetical protein Y47H9C.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 17-Mar-2000
C:Accession: T26972
R:Haris, B.
submitted to the EMBL Data Library, October 1998
A:Reference number: Z20293
A:Accession: T26972
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1111 <MIL>
A:Cross-references: EMBL:AL032657; PIDN:CAA21739.1; GSPDB:GN00019; CESP:Y47H9C.4
A:Experimental source: clone Y47H9C
C:Genetics:
A:Gene: CESP:Y47H9C.4
A:Map position: 1
A:introns: 50/2; 84/2; 150/1; 238/3; 342/3; 797/1; 851/1; 947/2; 1017/1; 1083/1
C:Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homology

Query Match 70.3%; Score 45; DB 2; Length 1111;
Best Local Similarity 70.0%; Pred. No. 8.5;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 CCDEVLOTKE 11
| | | | | | | | | |
Db 105 CCDEVLOTKE 114

RESULT 5

T05083
hypothetical protein T6K21.150 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cross)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-Jun-1999

C:Accession: T05083
R:Bevan, M.; Rieger, M.; Mueller-Auer, S.; Zipp, M.; Schaefer, M.; Bancroft, I.; Mewe submitted to the Protein Sequence Database, February 1998
A:Reference number: Z15397

A:Accession: T05083
A:Molecule type: DNA
A:Residues: 1-560 <BEV>
A:Cross-references: EMBL:AL021889
A:Experimental source: cultivar Columbia; BAC clone T6K21
C:Genetics:
A:Map position: 4
A:introns: 94/1; 141/1; 232/1; 269/3; 317/3
A:Note: T6K21.150

Query Match 64.1%; Score 41; DB 2; Length 560;
Best Local Similarity 63.6%; Pred. No. 23;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 ACCDEVLOTKE 11
| | | | | | | | | |
Db 232 ACDEVFEKE 242

RESULT 6

T07693
hypothetical protein F17N18.20 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cross)
C:Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 22-Oct-1999
C:Accession: T07693
R:Bevan, M.; Pohl, T.; Weizenecker, T.; Bancroft, I.; Mewe, H.W.; Mayer, K.F.X.; Sch submitted to the Protein Sequence Database, May 1999
A:Reference number: Z16098
A:Accession: T07693
A:Molecule type: DNA
A:Residues: 1-767 <BEV>
A:Cross-references: EMBL:AL049751; GSPDB:GN00062; ATSP:F17N18.20
A:Experimental source: cultivar Columbia; BAC clone F17N18
C:Genetics:
A:Gene: ATSP:F17N18.20
A:Map position: 4

Query Match 62.5%; Score 40; DB 2; Length 767;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 CCDEVYL 7
| | | | | | | | | |
Db 280 CCDEVYL 285

RESULT 7

A71676
hypothetical protein RP694 - Rickettsia prowazekii
C:Species: Rickettsia prowazekii
C:Date: 21-Nov-1998 #sequence_revision 21-Nov-1998 #text_change 03-Nov-2000
C:Accession: A71676
R:Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Slicheritz-Ponten, T.; Alsmark Nature 396, 133-140, 1998
A:Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.
A:Reference number: A71630; MUID:99039499; PMID:9823893
A:Accession: A71676
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-227 <AND>
A:Cross-references: GB:A7235272; GB:A7235269; NID:q3861033; PIDN:CA15131.1; PID:q386
A:Experimental source: strain Madrid E
C:Genetics:
A:Gene: RP694
C:Superfamily: conserved hypothetical protein H1191

Query Match 60.9%; Score 39; DB 2; Length 227;
Best Local Similarity 60.0%; Pred. No. 23;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 CCDEYLQTK 11
11:11:1
Db 136 CCPEYIOSFE 145

RESULT 8
T09635
C:Species: Lactobacillus delbrueckii
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C:Accession: T09635
R:Branny, P.; Delatorre, F.; Garel, J.R.
Microbiology 144, 905-914, 1998
A:Title: An operon coding for three glycolytic enzymes in Lactobacillus delbrueckii subs

A:Reference number: Z16788; MID:98240227; PMID:9579064
A:Accession: T09635
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-252 <BRA>
A:Cross-references: EMBL:AJ000339; NID:92624189; PIDN:CAA04016.1; PID:g2624193
A:Experimental source: subsp. bulgaricus, strain B107
C:Genetics:
A:Gene: tpi
C:Superfamily: triose-phosphate isomerase
C:Keywords: intramolecular oxidoreductase; isomerase

Query Match 59.4%; Score 38; DB 2; Length 252;
Best Local Similarity 60.0%; Pred. No. 37;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 CCDEYLQTK 11
11:11:1
Db 127 CCGESLEFRE 136

RESULT 9
T24089
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T24089
R:Matthews, L.
submitted to the EMBL Data Library, March 1996
A:Reference number: Z19839
A:Accession: T24089
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-1791 <WIL>
A:Cross-references: EMBL:Z70287; PIDN:CAA94301.1; GSPDB:GN00022; CESP:R09E10.7
A:Experimental source: clone R09E10
C:Genetics:
A:Gene: CESP:R09E10.7
A:Map position: 4
A:Intons: 78/2; 94/3; 250/2; 350/2; 674/3; 1079/2; 1155/2; 1450/3; 1593/3; 1642/2; 1690

Query Match 59.4%; Score 38; DB 2; Length 1791;
Best Local Similarity 71.4%; Pred. No. 2.1e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 2 CCDEYLQ 8
11:11:1
Db 1711 CCDEYILO 1717

RESULT 10
S28065
invH protein precursor - Salmonella typhimurium
C:Species: Salmonella typhimurium
C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 08-Oct-1999
C:Accession: S28065
R:Altmeier, R.M.; McNern, J.K.; Bossio, J.C.; Rosenshine, I.; Finlay, B.B.; Galan, J.E.

Mol. Microbiol. 7, 89-98, 1993
A:Title: Cloning and molecular characterization of a gene involved in Salmonella adhe
A:Reference number: S28064; MID:93172965; PMID:8382333

A:Accession: S28065
A:Molecule type: DNA
A:Residues: 1-147 <ALT>
A:Cross-references: EMBL:Z17242; NID:g47749; PIDN:CAA78942.1; PID:g47750
C:Genetics:
A:Gene: invH

Query Match 57.8%; Score 37; DB 2; Length 147;
Best Local Similarity 66.7%; Pred. No. 35;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 3 CCDEYLQTK 11
11:11:1
Db 85 CEKYKQTK 93

RESULT 11
S28064
invH protein precursor - Salmonella choleraesuis
C:Species: Salmonella choleraesuis
C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 08-Oct-1999
C:Accession: S28064
R:Altmeier, R.M.; McNern, J.K.; Bossio, J.C.; Rosenshine, I.; Finlay, B.B.; Galan, J.
Mol. Microbiol. 7, 89-98, 1993
A:Title: Cloning and molecular characterization of a gene involved in Salmonella adhe
A:Reference number: S28064; MID:93172965; PMID:8382333
A:Accession: S28064
A:Molecule type: DNA
A:Residues: 1-147 <ALT>
A:Cross-references: EMBL:Z17243; NID:g46862; PIDN:CAA78943.1; PID:g46863
C:Genetics:
A:Gene: invH

Query Match 57.8%; Score 37; DB 2; Length 147;
Best Local Similarity 66.7%; Pred. No. 35;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 3 CCDEYLQTK 11
11:11:1
Db 85 CEKYKQTK 93

RESULT 12
AH0852
cell adherence/invasion protein [imported] - Salmonella enterica subsp. enterica sero
C:Species: Salmonella enterica subsp. enterica serovar typh
A:Note: this species has also been called Salmonella typh
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 09-Nov-2001
C:Accession: AH0852
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Church
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Fair
, S.; Mout, S.; O'Gaora, P.
Nature 411, 848-852, 2001

A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens,
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica se
A:Reference number: AB0502; PMID:11677608
A:Accession: AH0852
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-147 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD06007.1; PID:g16503978; GSPDB:GN00176

C:Genetics:
A:Gene: invH

Query Match 57.8%; Score 37; DB 2; Length 147;
Best Local Similarity 66.7%; Pred. No. 35;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 3 CCDEYLQTK 11
11:11:1

Db 85 CEKYOTKE 93

RESULT 13

D64526
hypothetical protein HP0052 - Helicobacter pylori (strain 26695)

C:Species: Helicobacter pylori
C>Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 08-Oct-1999

C:Accession: D64526

R:omb, J.F.; White, O.; Kervatage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glock, A.; McInerney, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Matthey, L.; Nature 388, 539-547, 1997

A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.

A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.

A:Reference number: A64520; MUID:97394467; PMID:9252185

A:Accession: D64526

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-330 <TOM>

A:Cross-references: GB:AE000527; GB:AE000511; MID:g2313131; PIDN:AMD07127.1; PID:g2313131

C:Genetics:

A:Start codon: GTC

Query Match	57.8%;	Score 37;	DB 2;	Length 330;
Best Local Similarity	83.3%;	Pred. No. 70;		
Matches 5;	Conservative 1;	Mismatches 0;	Indels 0;	Gaps 0;

OY 1 ACCODEY 6

Db 282 ACCODY 287

RESULT 14

S62064

mitochondrial RNA splicing protein MRS2 - yeast (Saccharomyces cerevisiae)

N:Alternate names: protein O6247; protein YOR334W

C:Species: Saccharomyces cerevisiae

C>Date: 10-Apr-1996 #sequence_revision 19-Apr-1996 #text_change 20-Jun-2000

C:Accession: S62064; A42438; S27430; S67241; S71972

R:Parle, A.G.; Hand, N.J.; Goulding, S.G.; Wolfe, K.H.

submitted to the EMBL Data Library, June 1995

A:Description: Sequence of 29 kilobases around the PDR10 locus on the right arm of Sacch

A:Reference number: S62058

A:Accession: S62064

A:Molecule type: DNA

A:Residues: 1-470 <PAR>

A:Cross-references: EMBL:Z49821; MID:g1163062; PIDN:CAA89979.1; PID:g1163068

R:Wiesenberg, G.; Waldbert, M.; Schwen, R.J.

J. Biol. Chem. 267, 6963-6969, 1992

A:Title: The nuclear gene MRS2 is essential for the excision of group II introns from ye

A:Reference number: A42438; MUID:92202256; PMID:1551905

A:Accession: A42438

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-370, 'F', 372-447, 'D', 449-470 <WIE>

A:Cross-references: EMBL:M82916; MID:g171987; PIDN:AAA34795.1; PID:g171988

R:Goulding, S.E.; Hand, N.J.; Parle-McDermott, A.G.; Wolfe, K.H.

submitted to the Protein Sequence Database, July 1996

A:Reference number: S67233

A:Accession: S67241

A:Molecule type: DNA

A:Residues: 1-470 <GOU>

A:Cross-references: EMBL:Z75241; MID:g1420725; PIDN:CAA99656.1; PID:g1420727; MIPS:YOR33

A:Experimental source: strain S288C
R:Parle-McDermott, A.G.; Hand, N.J.; Goulding, S.E.; Wolfe, K.H.

Yeast 12, 999-1004, 1996

A:Title: Sequence of 29 kb around the PDR10 locus on the right arm of Saccharomyces c

A:Reference number: S71966; MUID:97051586; PMID:8896263

A:Accession: S71972

A>Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-470 <PAM>

A:Cross-references: EMBL:Z49821; MID:g1163062; PIDN:CAA89979.1; PID:g1163068

A>Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1995

C:Genetics:

A:Gene: SGD:MRS2

A:Cross-references: SGD:S0005861; MIPS:YOR334W

A:Map position: 15R

A:Genome: nuclear

C:Function:

A:Description: essential for the splicing of mitochondrial group II introns

C:Superfamily: mitochondrial RNA splicing protein MRS2

C:Keywords: mitochondrial; transmembrane protein

F:315-331/Domain: transmembrane #status predicted <TM1>

F:346-362/Domain: transmembrane #status predicted <TM2>

Query Match	57.8%;	Score 37;	DB 2;	Length 470;
Best Local Similarity	66.7%;	Pred. No. 95;		
Matches 6;	Conservative 1;	Mismatches 2;	Indels 0;	Gaps 0;

OY 3 CDEYLOTKKE 11

Db 276 CDEYVOQSE 284

RESULT 15

A42707

type II transposase homolog - Bacillus firmus (fragment)

C:Species: Bacillus firmus

C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 31-Dec-1993

C:Accession: A42707

R:Ivey, D.M.; Guffanti, A.A.; Shen, Z.; Kudyan, N.; Krulwich, T.A.

J. Bacteriol. 174, 4878-4884, 1992

A:Title: The cad gene product of alkaliphilic Bacillus firmus OF4 partially restores

A:Reference number: A42707; MUID:92332419; PMID:1321115

A:Accession: OF4

A:Contents: OF4

A:Accession: A42707

A:Molecule type: DNA

A:Residues: 1-840 <IVE>

A>Note: sequence extracted from NCBI backbone (NCBIN:108545, NCBIP:108546)

C:Superfamily: transposase Tn21

Query Match	57.8%;	Score 37;	DB 2;	Length 840;
Best Local Similarity	66.7%;	Pred. No. 1.6e+02;		
Matches 6;	Conservative 2;	Mismatches 1;	Indels 0;	Gaps 0;

OY 3 CDEYLOTKKE 11

Db 490 CDEYLVTKD 498

Search completed: November 18, 2002, 15:56:09

Job time : 15.129 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 18, 2002, 15:53:17 ; Search time 7.09677 Seconds
(without alignments)
64.288 Million cell updates/sec

Title: US-09-848-781-2

Perfect score: 64

Sequence: 1 ACCDEYIQTKK 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 segs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	54.4	84.4	289	PHLC_LISMO	P33378 listeria mo
2	38	59.4	252	TPIS_LACDE	O32757 lactobacill
3	38	59.4	367	HISB_PARDE	O51687 paracoccus
4	38	59.4	445	I12B_MOUSE	O91002 mus musculi
5	37	57.8	147	INVA_SALCH	P37422 salmonella
6	37	57.8	147	INVA_SALCH	P37423 salmonella
7	37	57.8	470	MRS2_YEAST	O01926 saccharomyc
8	37	57.8	1235	VGU2_CVMUH	P11225 murine coro
9	37	57.8	1376	VGU2_CVM4	P22432 murine coro
10	37	57.8	1376	VGU2_CVM4	P02385 murine coro
11	37	57.8	1407	T2D2_YEAST	P23255 saccharomyc
12	36	56.2	63	IBB_VICFA	P24651 vicia faba
13	36	56.2	72	IBB_PEA	P56679 pisum sativ
14	36	56.2	96	IBB_PEA	O41065 pisum sativ
15	36	56.2	114	IBB_PEA	O41066 pisum sativ
16	36	56.2	441	PEX3_YEAST	P28795 saccharomyc
17	36	56.2	538	VGLE_MOUSE	P09458 mumps virus
18	36	56.2	709	RRP2_INCIJ	P13828 influenza c
19	36	56.2	1324	VGU2_CVMA5	P11324 murine coro
20	36	56.2	2482	VWF_PIG	O28833 sus scrofa
21	36	56.2	2813	VWF_CANFA	O28833 sus scrofa
22	36	56.2	2813	VWF_CANFA	O28833 sus scrofa
23	35.5	55.5	1477	KELC_DROME	P04275 homo sapien
24	35	54.7	247	YOEM_BACSU	O04652 drosophila
25	35	54.7	296	MY88_HUMAN	P54458 bacillus su
26	35	54.7	296	MY88_HUMAN	O99936 homo sapien
27	35	54.7	377	FLRR_ECOS7	P22366 mus musculi
28	35	54.7	377	FLRR_ECOS7	O8x500 escherichia
29	35	54.7	377	FLRR_ECOS7	P37596 escherichia
30	35	54.7	377	FLRR_ECOS7	O8x500 escherichia
31	35	54.7	433	RTCL_DICDI	O82424 salmonella
32	35	54.7	513	YD17_SCHPO	O15746 dictyosteli
33	35	54.7	538	VGLE_MOUSE	O10440 schizosacch
					P11336 mumps virus

34	35	54.7	549	1	YB08_YEAST
35	35	54.7	951	1	LG84_HUMAN
36	35	54.7	951	1	LG84_RAT
37	35	54.7	1133	1	ATX9_TETTH
38	35	54.7	1172	1	AHM2_ARATH
39	35	54.7	1184	1	FBL2_HUMAN
40	35	54.7	1951	1	CIN3_HUMAN
41	35	54.7	1951	1	CIN3_RAT
42	34	53.1	61	1	IBB_ERYVA
43	34	53.1	117	1	SP41_HUMAN
44	34	53.1	117	1	SP41_MOUSE
45	34	53.1	251	1	TPIS_STRPY

ALIGNMENTS

RESULT 1	ID	PHLC_LISMO	STANDARD:	PRT:	289 AA.
AC	P33378:				
DT	01-FEB-1994 (Rel. 28, Created)				
DT	01-FEB-1994 (Rel. 28, Last sequence update)				
DT	15-JUN-2002 (Rel. 41, Last annotation update)				
DE	Phospholipase C precursor (PC 3.1.4.3) (PLC) (Phosphatidylcholine				
DE	cholephosphohydrolase) (lecithinase).				
GN	PLCB OR PRTC OR LMO0205.				
OS	Listeria monocytogenes.				
OC	Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.				
OX	NCBI_TaxID=1639;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=EGD / Serovar 1/2c;				
RX	MEDLINE=92258410; PubMed=1309513;				
RA	Vazquez-Boland J.-A., Kocks C., Dramsi S., Ohayon H., Geoffroy C.,				
RT	Mengaud J., Cossart P.;				
RT	"Nucleotide sequence of the lecithinase operon of Listeria				
RT	monocytogenes and possible role of lecithinase in cell-to-cell				
RT	spread";				
RL	Infect. Immun. 60:219-230(1992).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=EGD / Serovar 1/2a;				
RX	MEDLINE=21537279; PubMed=11679669;				
RA	Glaser P., Frangeul L., Buchrieser C., Rusnok C., Amend A.,				
RA	Baquerio F., Berche P., Biocker H., Brandt P., Chakraborty T.,				
RA	Charbit A., Chetoui F., Couve E., de Daruvar A., Dehoux P.,				
RA	Domin E., Dominguez-Bernal G., Duchaud E., Durant L., Dussauget O.,				
RA	Entian K.-D., Esnlt H., Garcia-del Portillo F., Garrido P.,				
RA	Gautier L., Goebel W., Gomez-Lopez N., Hain T., Haut J., Jackson D.,				
RA	Jones L.-M., Kaerst U., Kretz J., Kuhn M., Kunst F., Kurapkat G.,				
RA	Madueno E., Maitouram A., Mata Vicente J., Ng E., Nedjati H.,				
RA	Mordisco G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,				
RA	Rommel B., Rose M., Schueter T., Simoes N., Tierrez A.,				
RA	Vazquez-Boland J.-A., Voss H., Weiland J., Cossart P.;				
RT	"Comparative genomics of Listeria species.";				
RL	Science 294:849-852(2001).				
RP	SEQUENCE OF 144-152.				
RX	MEDLINE=91267617; PubMed=1904842;				
RA	Geoffroy C., Raveau J., Beretti J.L., Lecroisey A.,				
RA	Vazquez-Boland J.-A., Alouf J.E., Berche P.;				
RT	"Purification and characterization of an extracellular 29-kilodalton				

P38285	saccharomyc
O9bxd1	homo sapien
O9z2h4	rattus norv
O95050	tetrahymena
O64474	arabidopsis
P98095	homo sapien
O9ny46	homo sapien
P08104	rattus norv
P81705	erythrina v
O16550	homo sapien
O92199	mus musculi
P82478	streptococc

RT phospholipase C from *Listeria monocytogenes*.";
RL Infect. Immun. 59:2382-2388(1991).
CC -!- FUNCTION: IMPORTANT ROLE IN THE INFECTIOUS PROCESS. MAY
CC CONTRIBUTE TO EFFICIENT LYSIS OF THE TWO-MEMBRANE VACUOLES THAT
CC SURROUND THE BACTERIA AFTER DIRECT CELL-TO-CELL SPREAD.
CC
CC -!- CATALYTIC ACTIVITY: A phosphatidylcholine + H(2)O = 1,2-
CC diacylglycerol + choline phosphate.
CC
CC -!- COFACTOR: BINDS 3 ZINC IONS (By similarity).
CC
CC -!- SIMILARITY: STRONG, TO B.CEREUS PHOSPHOLIPASE C.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC
CC EMBL: M82881; AAA25270.1; -;
CC EMBL: X59723; CAA42408.1; -;
CC EMBL: AL591974; CAD00732.1; -;
CC PIR: C43868; C43868.
CC HSSP: P09598; 1AH7.
CC
CC Listlist: LMO00205; -;
CC InterPro: IPR001531; Zn_dep_PLPC.
CC Pfam: PF00882; Zn_dep_PLPC; 1.
CC PRINTS: PR00479; PRPHPLPASEC.
CC DR ProDom: PD003946; Zn_dep_PLPC; 1.
CC DR PROSITE: PS00384; PROKAR_ZN_DEPEND_PLPC; 1.
CC KM Hydroxase; Zinc; Zymogen; Signal; Complete proteome.
CC FT SIGNAL 1 25
CC FT PROPEP 26 51
CC FT CHAIN 52 289
CC FT METAL 52 289
CC FT METAL 65 65
CC FT METAL 106 106
CC FT METAL 120 120
CC FT METAL 169 169
CC FT METAL 173 173
CC FT METAL 179 179
CC FT METAL 193 193
CC FT METAL 197 197
CC FT CONFLICT 5 5
CC FT CONFLICT 13 13
CC FT CONFLICT 222 222
CC FT CONFLICT 222 222
CC SQ SEQUENCE 289 AA; 33277 MW; 0F35A2A3EDA6E372 CRC64;
Query Match 84.4%; Score 54; DB 1; Length 289;
Best Local Similarity 100.0%; Pred. No. 0.024;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACCDEYLTQT 9
DB 27 ACCDEYLTQT 35
RESULT 2
TPIS_LACDE
ID TPIS_LACDE STANDARD; PRT; 252 AA.
AC 032757;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Triosephosphate isomerase (EC 5.3.1.1) (TIM).
GN TPIS OR TPI.
OS Lactobacillus delbrueckii (subsp. bulgaricus).
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=1585;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BI07;
RC MEDLINE=98240227; PubMed=9579064;

RA Branny P., Delatorre F., Garel J.R.;
RT "An operon encoding three glycolytic enzymes in *Lactobacillus*
RT *delbrueckii* subsp. *bulgaricus*: glyceraldehyde-3-phosphate
RT dehydrogenase, phosphoglycerate kinase and triosephosphate
RT isomerase.";
RL Microbiology 144:905-914(1998).
CC -!- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate = glycero-
CC ne phosphate.
CC
CC -!- PATHWAY: PLAYS AN IMPORTANT ROLE IN SEVERAL METABOLIC PATHWAYS.
CC
CC -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC
CC -!- SIMILARITY: BELONGS TO THE TRIOSEPHOSPHATE ISOMERASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC
CC EMBL: AJ000339; CAA04016.1; -;
CC HSSP: P00943; 2BTM.
CC
CC InterPro: IPR000652; Triophos_ismrse.
CC Pfam: PF00121; TIM; 1.
CC DR ProDom: PD001005; Triophos_ismrse; 1.
CC DR TIGREMS: TIGR00419; tim; 1.
CC DR PROSITE: PS00171; TIM; 1.
CC KM Isomerase; Glycolysis; Gluconeogenesis; Fatty acid biosynthesis;
CC Pentose shunt.
CC FT ACT_SITE 96 96
CC FT ACT_SITE 168 168
CC FT ACT_SITE 168 168
CC SQ SEQUENCE 252 AA; 27383 MW; A39438DAA2745DAC CRC64;
Query Match 59.4%; Score 38; DB 1; Length 252;
Best Local Similarity 60.0%; Pred. No. 12;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 2 CCDEYLTQKE 11
DB 127 CCCESELTRE 136
RESULT 3
HIS8_PARDE
ID HIS8_PARDE STANDARD; PRT; 367 AA.
AC 051687;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Histidinol-phosphate aminotransferase (EC 2.6.1.9) (Imidazole aceto)-
DE phosphate transferase).
GN HIS8 OR HISH.
OS Paracoccus denitrificans.
OC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;
OC Paracoccus.
OX NCBI_TaxID=266;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Pd 1222;
RC MEDLINE=97195802; PubMed=9043133;
RA Page D., Pearce D.A., Norris H.A., Ferguson S.J.;
RT "The Paracoccus denitrificans ccmA, B and C genes: cloning and
RT sequencing, and analysis of the potential of their products to form a
RT haem or apo-c-type cytochrome transporter.";
RL Microbiology 143:563-576(1997).
CC -!- CATALYTIC ACTIVITY: L-histidinol-phosphate + 2-oxoglutarate - 3-
CC (imidazol-4-yl)-2-oxopropyl phosphate + L-glutamate.
CC
CC -!- COFACTOR: Pyridoxal phosphate (By similarity).
CC
CC -!- PATHWAY: Histidine biosynthesis; seventh step.
CC
CC -!- SIMILARITY: BELONGS TO CLASS-TI OF PYRIDOXAL-PHOSPHATE-DEPENDENT
CC AMINOTRANSFERASES. HISTIDINOL-PHOSPHATE AMINOTRANSFERASES
CC SUBFAMILY.
CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
CC EMBL: Z71971; CAA96498.1; -
CC InterPro: IPR004839; AminoTransf1/2.
CC Pfam: PF00155; aminotran_1.2; 1.
CC TIGRfam: TIGR01141; hisc_1.
CC PROSITE: PS00599; AA_TRANSFERR_CLASS_2; FALSE_NEG.
CC Histidine biosynthesis; Transferase; AminoTransferase;
CC Pyridoxal phosphate.
CC BINDING 221 221 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
CC SEQUENCE 367 AA; 39501 MW; 74B755668FED2338F CRC64;

Query Match 59.4%; Score 38; DB 1; Length 367;
Best Local Similarity 75.0%; Pred. No. 18;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 3 CDEYLOTKE 10
Db 316 CDEYLOTKE 323

RESULT 4
ID I12B_MOUSE STANDARD; PRT; 445 AA.
AC 09R002;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Interferon-activatable protein 202b (I1i-202b) (Interferon-inducible
DE protein p202b).
GN I1I202B.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/Sv;
RX MEDLINE=99425269; PubMed=10493828;
RA Wang H., Chatterjee G., Meyer J.J., Liu C.J., Manjunath N.A.,
RA Bray-Ward P., Lengyel P.,
RT "Characteristics of three homologous 202 genes (I1i202a, I1i202b, and
RT I1i202c) from the murine interferon-activatable gene 200 cluster.";
RL Genomics 60:281-294(1999).
CC -I- FUNCTION: INHIBITS THE TRANSCRIPTIONAL ACTIVITY OF SEVERAL
CC TRANSCRIPTION FACTORS.
CC -I- SUBUNIT: MAY BIND TO SEVERAL TRANSCRIPTION FACTORS.
CC -I- SUBCELLULAR LOCATION: ACCUMULATES FIRST IN THE CYTOPLASM, AND IS
CC CHROMATIN-ASSOCIATED (BY SIMILARITY).
CC -I- INDUCTION: By beta interferon; strongly.
CC -I- SIMILARITY: CONTAINS 2 HIN-200 DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
CC EMBL: AF140672; AAF04260.1; -
CC MCD: MGI:1347083; I1i202b.
CC InterPro: IPR004021; HIN_dom.
CC Pfam: PF02760; HIN; 2.
CC PROSITE: PS50834; HIN_200; 2.
CC Interferon induction; Nuclear protein; Repeat; Phosphorylation.

FT DOMAIN 46 243 HIN-200 1.
FT DOMAIN 244 441 HIN-200 2.
FT DOMAIN 82 288 2 x 8 AA REPEATS OF M-F-H-A-T-V-A-T.
FT REPEAT 82 89 1.
FT REPEAT 281 288 2.
FT DOMAIN 343 346 POLY-SER.
FT SITE 84 84 MEDIATES INTERACTION WITH 53BP1 (BY
FT SITE SIMILARITY).
FT SITE 283 283 MEDIATES INTERACTION WITH 53BP1 (BY
FT SITE SIMILARITY).
SQ SEQUENCE 445 AA; 50395 MW; AF049472632A7720 CRC64;

Query Match 59.4%; Score 38; DB 1; Length 445;
Best Local Similarity 77.8%; Pred. No. 21;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 3 CDEYLOTKE 11
Db 357 CDEYLOTKE 365

RESULT 5
ID INVH_SALICH STANDARD; PRT; 147 AA.
AC P37422;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-FEB-1995 (Rel. 31, Last annotation update)
DE Invasion protein invh precursor.
GN INVH.
OS Salmonella cholerae-suis (Salmonella enterica).
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=591;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SL2824;
RX MEDLINE=93172965; PubMed=8382333;
RA Altmeyer R.M., McKern J.K., Bossio J.C., Rosenshine I.,
RA Finlay B.B., Galan J.E.,
RT "Cloning and molecular characterization of a gene involved in
RT Salmonella adherence and invasion of cultured epithelial cells.";
RL Mol. Microbiol. 7:89-98(1993).
CC -I- FUNCTION: NECESSARY FOR EFFICIENT ADHERENCE AND ENTRY OF THE
CC ORGANISMS INTO CULTURED EPITHELIAL CELLS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
CC EMBL: Z17243; CAA78943.1; -
CC DR PIR: S28064; S28064.
CC KW Virulence; Signal.
FT SIGNAL 1 28 POTENTIAL.
FT CHAIN 29 147 INVASION PROTEIN INVH.
SQ SEQUENCE 147 AA; 16463 MW; FEF5079DBE8B3762 CRC64;

Query Match 57.8%; Score 37; DB 1; Length 147;
Best Local Similarity 66.7%; Pred. No. 11;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 3 CDEYLOTKE 11
Db 85 CDEYLOTKE 93

RESULT 6
ID INVH_SALICH STANDARD; PRT; 147 AA.
ID INVH_SALICH STANDARD; PRT; 147 AA.

```

AC      P37423;
AD      01-OCT-1994 (Rel. 30, Created)
DT      01-OCT-1994 (Rel. 30, Last sequence update)
DT      15-JUN-2002 (Rel. 41, Last annotation update)
DE      Invasion protein InvH precursor.
GN      INVH OR STM2900.
OS      Salmonella typhimurium.
OC      Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC      Salmonella.
OX      NCBI_TaxID=602;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN-SRI / SL1344;
RX      MEDLINE=93172965; PubMed=8382333;
RA      Altmeier R.M., McNeir J.K., Bossio J.C., Rosenshine I.,
RA      Finlay B.B., Galan J.E.;
RT      "Cloning and molecular characterization of a gene involved in
RT      Salmonella adherence and invasion of cultured epithelial cells.";
RL      Mol. Microbiol. 7:89-98(1993).
RN      [2]
RP      SEQUENCE FROM N.A.
RC      STRAIN-S4194;
RX      MEDLINE=97221599; PubMed=9068645;
RA      Boyd E.F., Li J., Ochman H., Selander R.K.;
RT      "Comparative genetics of the inv-spa invasion gene complex of
RT      Salmonella enterica.";
RL      J. Bacteriol. 179:1985-1991(1997).
RN      [3]
RP      SEQUENCE FROM N.A.
RC      STRAIN-LT2 / SGSC412 / ATCC 700720;
RX      MEDLINE=21534948; PubMed=11677609;
RA      McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA      Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA      Leonard S., Nguyen C., Scott K., Holmes A., Grewal A., Mufvany E.,
RA      Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA      Waterston R., Wilson R.K.;
RT      "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT      LT2.";
RL      Nature 413:852-856(2001).
CC      -!- FUNCTION: NECESSARY FOR EFFICIENT ADHERENCE AND ENTRY OF THE
CC      ORGANISMS INTO CULTURED EPITHELIAL CELLS.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; Z17242; CAAT8942.1; -
DR      EMBL; U84273; AAC45070.1; -
DR      EMBL; AE008833; AAL21780.1; -
DR      PIR; S28065; S28065.
DR      StyGene; SG10189; INVH.
KW      Virulence; Signal; Complete proteome.
FT      SIGNAL 1 28
FT      CHAIN 29 147
FT      SEQUENCE 147 AA; 16455 MW; A02A407F1BC6DCFC CXC64;
OY      3 CDEVLQTKK 11
OY      1::1 1111
DB      85 CEKYNQTKK 93

RESULT 7
MRS2_YEAST STANDARD: PRT: 470 AA.
ID MRS2_YEAST
RC 001926;

```

```

DT      01-JUL-1993 (Rel. 26, Created)
DT      01-OCT-1996 (Rel. 34, Last sequence update)
DT      15-JUN-2002 (Rel. 41, Last annotation update)
DE      RNA splicing protein MRS2, mitochondrial precursor.
DE      MRS2 OR YOR334W.
OS      Saccharomyces cerevisiae (Baker's Yeast).
OC      Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC      Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX      NCBI_TaxID=4932;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=DRY747;
RX      MEDLINE=92202256; PubMed=1551905;
RA      Wiesenberger G., Walther M., Scheyen R.J.;
RT      "The nuclear gene MRS2 is essential for the excision of group II
RT      introns from yeast mitochondrial transcripts in vivo.";
RL      J. Biol. Chem. 267:6963-6969(1992).
RN      [2]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=94073984; PubMed=8252639;
RA      Walther M., Ragnini A., Jank B., Teply R., Wiesenberger G.,
RA      Scheyen R.J.;
RT      "A multitude of suppressors of group II intron-splicing defects in
RT      yeast.";
RL      Curr. Genet. 24:301-306(1993).
RN      [3]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=97051586; PubMed=8896263;
RA      Parle-mcdermott A.G., Hand N.J., Goulding S.G., Wolfe K.H.;
RT      "Sequence of 29 kb around the PDR10 locus on the right arm of
RT      Saccharomyces cerevisiae chromosome XV: similarity to part of
RT      chromosome I.";
RL      Yeast 12:999-1004(1996).
CC      -I: FUNCTION: MRS2 IS ESSENTIAL FOR THE SPLICING OF MRNA GROUP II
CC      INTRONS AND ITS ABSENCE MAY DISTURB THE ASSEMBLY OF MITOCHONDRIAL
CC      MEMBRANE COMPLEXES. IT ALSO SUPPRESSES A VARIETY OF MITOCHONDRIAL
CC      INTRON MUTATIONS.
CC      -I: SUBCELLULAR LOCATION: Mitochondrial.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL collaboration
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; M82916; AAA34795.1; -
DR      EMBL; Z49821; CA89979.1; -
DR      EMBL; Z75241; CA99656.1; -
DR      PIR; S27430; S27430.
DR      PIR; A42438; A42438.
DR      SGD; S0005861; MRS2.
KW      mRNA processing; Mitochondrion; Transit peptide.
FT      TRANSLAT 1
FT      CHAIN ? 470
FT      CONFLICT 371 371 RNA SPLICING PROTEIN MRS2.
FT      CONFLICT 448 448 S -> F (IN REF. 1 AND 2).
FT      FT E -> D (IN REF. 1 AND 2).
SQ      SEQUENCE 470 AA; 54202 MW; EPEBA11765A8B8B31 CRC64;

Query Match 57.8%; Score 37; DB 1; Length 470;
Best local similarity 66.7%; Pred. No. 33;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY      3 CDEYLTQKE 11
DB      276 CDEYQOSE 284

RESULT 8
ID      VGL2_CVMDH STANDARD; PRT; 1235 AA.
AC      P11225;

```

DT 01-JUL-1989 (rel. 11, Created)
DT 01-JUL-1989 (rel. 11, Last sequence update)
DT 15-JUN-2002 (rel. 41, Last annotation update)
DE E2 glycoprotein precursor (Spike glycoprotein) (Peplomer protein)
DE [Contains: Spike protein S1 (90B); Spike protein S2 (90A)].
GN S.
OS Murine coronavirus MHV (strain JHM).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC Coronaviridae; Coronavirus.
OX NCBI_TaxID=11144;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87111467; PubMed=3027248;
RA Schmidt I., Skinner M.A., Siddeil S.G.;
RT "Nucleotide sequence of the gene encoding the surface projection
glycoprotein of coronavirus MHV-JHM.";
RT J. Gen. Virol. 68:47-56(1987).
CC -I- FUNCTION: THE PEPLIMER PROTEIN MEDIATES THE BINDING OF VIRIONS
TO THE HOST CELL RECEPTOR AND IS INVOLVED IN MEMBRANE FUSION
AND IN SYNCYTUM FORMATION.
CC -I- SUBCELLULAR LOCATION: Type I membrane protein.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC EMBL: X04797; CAA28484.1; -.
CC PIR: A33095; VGJHMJ.
CC InterPro: IPR002552; Corona_S2.
DR Pfam: PF01601; Corona_S2; 1.
KW Glycoprotein; Envelope protein; Transmembrane; Signal.
FT SIGNAL 1 10
FT CHAIN 1 1235
FT CHAIN 1 1235
FT CHAIN 1 1235
FT CHAIN 629 1235
FT DOMAIN 1 1174
FT TRANSMEM 1175 1197
FT DOMAIN 1198 1235
FT DOMAIN 1198 1235
FT CARBOHYD 31 31
FT CARBOHYD 60 60
FT CARBOHYD 134 134
FT CARBOHYD 192 192
FT CARBOHYD 357 357
FT CARBOHYD 435 435
FT CARBOHYD 536 536
FT CARBOHYD 568 568
FT CARBOHYD 576 576
FT CARBOHYD 599 599
FT CARBOHYD 648 648
FT CARBOHYD 665 665
FT CARBOHYD 804 804
FT CARBOHYD 1091 1091
FT CARBOHYD 1101 1101
FT CARBOHYD 1120 1120
FT CARBOHYD 1136 1136
FT CARBOHYD 1157 1157
SQ SEQUENCE 1235 AA; 136653 MW; 25962AD6C1F92DD2 CRC64;
Query Match 57.8%; Score 37; DB 1; Length 1235;
Best Local Similarity Pred. 83.3%; Mismatches 0; Indels 0; Gaps 0;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

VG12_CVMA ID VG12_CVMA STANDARD; PRT: 1376 AA.
AC P22432;
DT 01-AUG-1991 (rel. 19, Created)
DT 01-AUG-1991 (rel. 19, Last sequence update)
DT 16-OCT-2001 (rel. 40, Last annotation update)
DE E2 glycoprotein precursor (Spike glycoprotein) (Peplomer protein)
DE [Contains: Spike protein S1 (90B); Spike protein S2 (90A)].
GN S.
OS Murine coronavirus MHV (strain wild type 4) (MHV-4).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC Coronaviridae; Coronavirus.
OX NCBI_TaxID=12760;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91353381; PubMed=1966429;
RA Parker S.E., Buchmeier M.J.;
RT "RNA sequence analysis of the E2 genes of wildtype and
neuroattenuated mutants of MHV-4 reveals a hypervariable domain.";
RT Adv. Exp. Med. Biol. 276:395-402(1990).
CC -I- FUNCTION: THE PEPLIMER PROTEIN MEDIATES THE BINDING OF VIRIONS
TO THE HOST CELL RECEPTOR AND IS INVOLVED IN MEMBRANE FUSION
AND IN SYNCYTUM FORMATION.
CC -I- SUBCELLULAR LOCATION: Type I membrane protein.
CC -I- SIMILARITY: NEARLY IDENTICAL TO THE E2 GLYCOPROTEINS FROM MHV-JHM
AND MHV-A59 STRAINS, EXCEPT FOR AN N-TERMINAL INSERTION.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC EMBL: M32789; AAA46456.1; -.
CC EMBL: S51114; AAB19590.1; -.
CC PIR: A33748; VGJHMJ2.
CC InterPro: IPR002552; Corona_S2.
DR Pfam: PF01601; Corona_S2; 1.
KW Glycoprotein; Envelope protein; Transmembrane; Signal.
FT SIGNAL 1 14
FT CHAIN 1 1376
FT CHAIN 1 1376
FT CHAIN 770 1376
FT CHAIN 15 1376
FT DOMAIN 15 1320
FT TRANSMEM 1321 1338
FT DOMAIN 1339 1376
FT DOMAIN 1339 1356
FT DOMAIN 429 599
FT CARBOHYD 31 31
FT CARBOHYD 60 60
FT CARBOHYD 134 134
FT CARBOHYD 192 192
FT CARBOHYD 357 357
FT CARBOHYD 435 435
FT CARBOHYD 582 582
FT CARBOHYD 677 677
FT CARBOHYD 709 709
FT CARBOHYD 717 717
FT CARBOHYD 740 740
FT CARBOHYD 789 789
FT CARBOHYD 806 806
FT CARBOHYD 945 945
FT CARBOHYD 1232 1232
SQ SEQUENCE 1376 AA; 137653 MW; 25962AD6C1F92DD2 CRC64;
Query Match 57.8%; Score 37; DB 1; Length 1376;
Best Local Similarity Pred. 83.3%; Mismatches 0; Indels 0; Gaps 0;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

FT CARBOHYD 1242 1242 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1261 1261 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1277 1277 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1298 1298 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1376 AA; 151882 MW; 88C01B97B252094E CRC64;

Query Match 57.8%; Score 37; DB 1; Length 1376;
Best Local Similarity 83.3%; Pred. No. 91;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCDEX 6
Db 1354 SCCDEX 1359

RESULT 10
VGL2_CVMJC
ID VGL2_CVMJC STANDARD; PRT; 1376 AA.
AC 002385;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE E2 glycoprotein precursor (spike glycoprotein) (Peplomer protein)
DE (Contains: Spike protein S1 (90B); Spike protein S2 (90A)).
GN S.
OS Murine coronavirus MHV (strain JHMV / variant CL-2).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC Coronaviridae; Coronavirus.
OX NCBI_TaxID=33735;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-92268864; PubMed-1316938;
RA Taguchi F., Ikeda T., Shida H.;
RT "Molecular cloning and expression of a spike protein of neurovirulent
RT murine coronavirus JHMV variant CL-2."
RL J. Gen. Virol. 73:1065-1072(1992).
CC -1- FUNCTION: THE PEPLIMER PROTEIN MEDIATES THE BINDING OF VIRIONS
CC TO THE HOST CELL RECEPTOR AND IS INVOLVED IN MEMBRANE FUSION
CC AND IN SYNCYTUM FORMATION.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D10235; BAA01085.1; -
DR PIR: J01534; J01534.
DR InterPro: IPR002552; Corona_S2.
DR Pfam: PF01601; Corona_S2; 1.
KW Glycoprotein; Envelope protein; Transmembrane; Signal.
FT SIGNAL 1 14
FT CHAIN 15 1376 E2 GLYCOPROTEIN.
FT CHAIN 15 769 SPIKE PROTEIN S1.
FT CHAIN 770 1376 SPIKE PROTEIN S2.
FT DOMAIN 15 1320 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1321 1338 POTENTIAL.
FT DOMAIN 1339 1376 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 1339 1356 CYS-RICH.
FT CARBOHYD 31 31 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 60 60 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 134 134 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 192 192 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 357 357 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 435 435 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 677 677 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 709 709 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 717 717 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 789 789 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 806 806 N-LINKED (GLCNAC. . .) (POTENTIAL).

```

```

FT CARBOHYD 945 945 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1232 1232 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1242 1242 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1261 1261 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1277 1277 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1298 1298 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1376 AA; 152041 MW; 98C30DD979F9E75 CRC64;

Query Match 57.8%; Score 37; DB 1; Length 1376;
Best Local Similarity 83.3%; Pred. No. 91;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCDEX 6
Db 1354 SCCDEX 1359

RESULT 11
T2D2_YEAST
ID T2D2_YEAST STANDARD; PRT; 1407 AA.
AC P23255;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Transcription initiation factor TFIID 150 kDa subunit (TAFII-150)
DE (TSM1 protein) (TSM-1).
GN TSM1 OR TAFI150 OR YCR042C OR YCR42C OR YCR724.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-92160397; PubMed-1789011;
RA Jaquet M., Bunler J.-M., Idorria F., Francinques-Galliard M.-C.,
RA Soustelle C.;
RT "The MAT locus revisited within a 9.8 kb fragment of chromosome III
RT containing BUD5 and two new open reading frames."
RL Yeast 7:881-888(1991).
RN [2]
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN-RY570;
RX MEDLINE-92035068; PubMed-1840512;
RA Ray B.L., White C.I., Haber J.E.;
RT "The TSM1 gene of Saccharomyces cerevisiae overlaps the MAT locus."
RL Curr. Genet. 20:25-31(1991).
RN [3]
RP SEQUENCE OF 165-174 AND 991-1000, AND CHARACTERIZATION.
RC STRAIN-YPH252;
RX MEDLINE-95396770; PubMed-7667272;
RA Poon D., Bai Y., Campbell A.M., Bjorklund S., Kim Y.-J., Zhou S.,
RA Kornberg R.D., Well P.A.;
RT "Identification and characterization of a TFIID-like multiprotein
RT complex from Saccharomyces cerevisiae."
RL Proc. Natl. Acad. Sci. U.S.A. 92:8224-8228(1995).
CC -1- FUNCTION: TAFs ARE COMPONENTS OF THE TRANSCRIPTION FACTOR IID
CC (TFIID) COMPLEX THAT ARE ESSENTIAL FOR MEDIATING REGULATION OF
CC RNA POLYMERASE TRANSCRIPTION.
CC -1- SUBUNIT: TFIID IS COMPOSED OF TATA BINDING PROTEIN (TBP) AND A
CC NUMBER OF TBP-ASSOCIATED FACTORS (TAFs) WHOSE MW RANGE FROM
CC 25-150 kDa.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: NO DROSOPHILA TAFII-150.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X63853; CAA45337.1; -

```

DR EMBL; M60486; AAA35179.1; -
DR EMBL; X59720; CAA42290.1; -
DR PIR; S19455; BMBYML.
DR MEROPS; M01.UNM; -
DR TRANSFAC; T01269; -
DR SGD; S0000638; TSM1.
KW Transcription regulation; Nuclear protein.
FT DOMAIN 259 334 ASP/GLU-RICH (ACIDIC).
FT DOMAIN 1285 1350 HIGHLY CHARGED.
FT CONFLICT 323 323 E -> G (TN REF. 2).
SQ SEQUENCE 1407 AA; 161470 MW; 184C8ED735443F8 CRC64;

Query Match 57.8%; Score 37; DB 1; Length 1407;
Best Local Similarity 60.0%; Pred. No. 93;
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 2 CCDEYLOTK 11
DB 346 CCSEYSNIKE 355

RESULT 12
IBB_VICFA STANDARD; PRT; 63 AA.
AC P24661;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE Bowman-Birk type proteinase inhibitor (FBI).
OS Vicia faba (Broad bean).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciaeae; Vicia.
OX NCBI_TaxID=3906;
RN [1]
RP SEQUENCE.
RC TISSUE=Seed;
RA MEDLINE-92176172; PubMed-1794984;
RA Masao T., Inai F., Tsuji I., Tashiro M., Iwami K., Ibuki F.;
RT "The amino acid sequence of a Bowman-Birk type proteinase inhibitor
RT from faba beans (Vicia faba L.)."
RL J. Biochem. 110:951-955(1991).
CC -1- FUNCTION: THIS INHIBITOR HAS TWO DOMAINS, EACH WITH SEPARATE
CC AMINOPEPTIDASE ACTIVITY. INHIBITS BOVINE TRYPSIN AND CHYMOTRYPSIN,
CC IN A MOLAR RATIO OF 1:1. THE TRYPSIN INHIBITION OF FBI IS
CC INDEPENDENT OF CHYMOTRYPSIN INHIBITION, BUT THE CHYMOTRYPSIN
CC INHIBITION IS NOT COMPLETELY INDEPENDENT OF TRYPSIN INHIBITION.
CC -1- SIMILARITY: BELONGS TO THE BOWMAN-BIRK SERINE PROTEASE INHIBITORS
CC FAMILY.
DR PIR; JX0198; TIVE.
DR HSSP; P56679; IPBI.
DR InterPro; IPR000877; Bowman-Birk_leg.
DR Pfam; PF00228; Bowman-Birk_leg; 1.
DR ProDom; PD002168; Bowman-Birk_leg; 1.
DR SMART; SM00269; BowB; 1.
DR PROSITE; PS00281; BOWMAN_BIRK; 1.
KW Serine protease inhibitor.
FT ACT_SITE 16 17 INTERACTION WITH TRYPSIN.
FT ACT_SITE 42 43 INTERACTION WITH CHYMOTRYPSIN.
FT DISULFID 8 61 BY SIMILARITY.
FT DISULFID 9 24 BY SIMILARITY.
FT DISULFID 12 57 BY SIMILARITY.
FT DISULFID 14 22 BY SIMILARITY.
FT DISULFID 31 38 BY SIMILARITY.
FT DISULFID 35 50 BY SIMILARITY.
FT DISULFID 40 48 BY SIMILARITY.
SQ SEQUENCE 63 AA; 6989 MW; 9FDEB47973C14B34 CRC64;

Query Match 56.2%; Score 36; DB 1; Length 63;
Best Local Similarity 70.0%; Pred. No. 7.3;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 ACCDEYLOTK 10

DB 7 ACCDTOLCTR 16
IIII I II
RESULT 13
IBB_PEA STANDARD; PRT; 72 AA.
AC P56679;
DT 15-JUL-1999 (Rel. 38, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Seed trypsin/chymotrypsin inhibitor IIV (PST-IIV) [Contains: Seed
DE trypsin/chymotrypsin inhibitor II (PST-II)].
OS Pisum sativum (Garden pea).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciaeae; Pisum.
OX NCBI_TaxID=3888;
RN [1]
RP SEQUENCE.
RC STRAIN=cv. Filienne; TISSUE=Seed;
RA MEDLINE-97299941; PubMed-9155090;
RA Quillien L., Ferrasson E., Molle D., Gueguen J.;
RT "Trypsin inhibitor polymorphism: multigene family expression and
RT posttranslational modification."
RL J. Protein Chem. 16:195-203(1997).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS).
RC TISSUE=Seed;
RA MEDLINE-99141373; PubMed-9887273;
RA Li de la Sierra L., Quillien L., Flecker P., Gueguen J., Brunle S.;
RT "Dimeric crystal structure of a Bowman-Birk protease inhibitor from
RT pea seeds."
RL J. Mol. Biol. 285:1195-1207(1999).
CC -1- FUNCTION: INHIBITOR OF TRYPSIN AND OF CHYMOTRYPSIN. MAY FUNCTION
CC AS A NATURAL PHYTOCHEMICAL DEFENSE AGAINST PREDATORS.
CC -1- TISSUE SPECIFICITY: SEED.
CC -1- DEVELOPMENTAL STAGE: DURING DESICCATION STAGE OF SEED DEVELOPMENT
CC INCREASING ACTIVITY SEEMS TO BE ASSOCIATED WITH APPEARANCE OF
CC ISOFORM (PST II) WHICH HAS A STRONGER AFFINITY FOR TRYPSIN.
CC -1- SIMILARITY: BELONGS TO THE BOWMAN-BIRK SERINE PROTEASE INHIBITORS
CC FAMILY.
DR PDB; 1PBI; 16-FEB-99.
DR InterPro; IPR000877; Bowman-Birk_leg.
DR Pfam; PF00228; Bowman-Birk_leg; 1.
DR ProDom; PD002168; Bowman-Birk_leg; 1.
DR SMART; SM00269; BowB; 1.
DR PROSITE; PS00281; BOWMAN_BIRK; FALSE NEG.
KW Serine protease inhibitor; 3D-structure; Multigene family.
FT CHAIN 1 72 SEED TRYPSIN/CHYMOTRYPSIN INHIBITOR IIV.
FT PROPEP 64 72 REMOVED IN PST II FORM.
FT ACT_SITE 42 43 INTERACTION WITH TRYPSIN.
FT DISULFID 8 61 INTERACTION WITH CHYMOTRYPSIN.
FT DISULFID 9 24
FT DISULFID 12 57
FT DISULFID 14 22
FT DISULFID 31 38
FT DISULFID 35 50
FT DISULFID 40 48
SQ SEQUENCE 72 AA; 7864 MW; F12E71FAAD9A4BF7 CRC64;

Query Match 56.2%; Score 36; DB 1; Length 72;
Best Local Similarity 70.0%; Pred. No. 8.3;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 ACCDEYLOTK 10
DB 7 ACCDTOLCTR 16

RESULT 14

IBBA_PEA
ID IBBA_PEA STANDARD; PRT; 96 AA.
AC Q41065;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Seed trypsin/chymotrypsin inhibitor IVA precursor (PSTI IVA) (T112-36)
DE [contains: Seed trypsin/chymotrypsin inhibitor I (PSTI I)] (Fragment).
GN T11236.
OS Pisum sativum (Garden pea).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciae; Pisum.
OC NCBI_TaxID=3888;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 25-56.
RC STRAIN-CV. Birtle, and cv. J1 2;
RX MEDLINE=95180408; PubMed=7875292;
RA Domoney C., Welham T., Sidebottom C., Firmin J.L.;
RT "Multiple isoforms of Pisum trypsin inhibitors result from
modification of two primary gene products.";
RL FEBS Lett. 360:15-20(1995).
RN [2]
RP SEQUENCE OF 25-96.
RC STRAIN-CV. Friene, TISSUE=Seed;
RX MEDLINE=96109762; PubMed=8593187;
RA Ferrasson E., Quillien L., Gueguen J.;
RT "Amino acid sequence of a Bowman-Birk proteinase inhibitor from pea
seeds.";
RL J. Protein Chem. 14:467-475(1995).
RN [3]
RP PROCESSING.
RC STRAIN-CV. Friene, TISSUE=Seed;
RX MEDLINE=97299941; PubMed=9155090;
RA Quillien L., Ferrasson E., Mollie D., Gueguen J.;
RT "Trypsin inhibitor polymorphism: multigene family expression and
posttranslational modification.";
RL J. Protein Chem. 16:195-203(1997).
CC -1- FUNCTION: INHIBITOR OF TRYPSIN AND OF CHYMOTRYPSIN. MAY FUNCTION
AS A NATURAL PHYTOCHEMICAL DEFENSE AGAINST PREDATORS.
CC -1- TISSUE SPECIFICITY: SEED.
CC -1- DEVELOPMENTAL STAGE: DURING DESICCATION STAGE OF SEED DEVELOPMENT
INCREASING ACTIVITY SEEMS TO BE ASSOCIATED WITH APPEARANCE OF
ISOFORM I (PSTI I) WHICH HAS A STRONGER AFFINITY FOR TRYPSIN.
CC -1- SIMILARITY: BELONGS TO THE BOWMAN-BIRK SERINE PROTEASE INHIBITORS
FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation-
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X83211; CAA58213.1; -.
DR HSSP: P56679; IPBI.
DR InterPro: IPR000877; Bowman-Birk_Leg.
DR Pfam: PF00228; Bowman-Birk_Leg; 1.
DR ProDom: PD002168; Bowman-Birk_Leg; 1.
DR SMART: SM00269; BowB; 1.
DR PROSITE: PS00281; BOWMAN_BIRK; 1.
KW Serine protease inhibitor; Signal; Multigene family.
FT NON_TER 1 1
FT SIGNAL 1 10
FT PROPEP 11 24
FT CHAIN 25 96
FT CHAIN 25 96
FT PROPEP 88 96
FT ACT_SITE 40 41
FT ACT_SITE 66 67
FT DISULFID 32 85
BY SIMILARITY.

FT DISULFID 33 48
FT DISULFID 36 81
FT DISULFID 38 46
FT DISULFID 55 62
FT DISULFID 59 74
FT DISULFID 64 72
FT CONFLICT 52 52
SQ SEQUENCE 96 AA; 10544 MW; 654F0E0069BF2 CRC64;
Query Match 56.2%; Score 36; DB 1; Length 96;
Best Local Similarity 70.0%; Pred. No. 11;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 ACCDEYLOTK 10
Db 31 ACCDTCLCTK 40
RESULT 15
IBB2_PEA
ID IBB2_PEA STANDARD; PRT; 114 AA.
AC Q41066;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Seed trypsin/chymotrypsin inhibitor T15-72 precursor.
GN T1572.
OS Pisum sativum (Garden pea).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciae; Pisum.
OC NCBI_TaxID=3888;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CV. Birtle;
RX MEDLINE=95180408; PubMed=7875292;
RA Domoney C., Welham T., Sidebottom C., Firmin J.L.;
RT "Multiple isoforms of Pisum trypsin inhibitors result from
modification of two primary gene products.";
RL FEBS Lett. 360:15-20(1995).
CC -1- FUNCTION: INHIBITOR OF TRYPSIN AND OF CHYMOTRYPSIN. MAY FUNCTION
AS A NATURAL PHYTOCHEMICAL DEFENSE AGAINST PREDATORS.
CC -1- TISSUE SPECIFICITY: SEED.
CC -1- SIMILARITY: BELONGS TO THE BOWMAN-BIRK SERINE PROTEASE INHIBITORS
FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation-
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X83210; CAA58212.1; -.
DR HSSP: P56679; IPBI.
DR InterPro: IPR000877; Bowman-Birk_Leg.
DR Pfam: PF00228; Bowman-Birk_Leg; 1.
DR ProDom: PD002168; Bowman-Birk_Leg; 1.
DR SMART: SM00269; BowB; 1.
DR PROSITE: PS00281; BOWMAN_BIRK; 1.
KW Serine protease inhibitor; Signal; Multigene family.
FT SIGNAL 1 28
FT PROPEP 29 42
FT CHAIN 43 114
FT ACT_SITE 58 59
FT ACT_SITE 84 85
FT DISULFID 50 103
FT DISULFID 51 66
FT DISULFID 54 99
FT DISULFID 56 64
BY SIMILARITY.

FT	DISULFID	73	80	BY SIMILARITY.
FT	DISULFID	77	92	BY SIMILARITY.
FT	DISULFID	82	90	BY SIMILARITY.
SO	SEQUENCE	114 AA;	12597 MW;	B60FC0768E8AB2A3 CRC64;

Query Match 56.28; Score 36; DB 1; Length 114;
 Best Local Similarity 70.08; Pred. No. 13;
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY	1	ACDDEVLOTK	10
		1111	111
Db	49	ACCDTCLCTK	58

Search completed: November 18, 2002, 15:53:48
 Job time : 8.09677 secs

THIS PAGE BLANK (USPTO)

AC 08T3A7;
DT 01-JUN-2002 (TReMBLrel. 21, Created)
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Y47H9C.4b protein.
GN Y47H9C.4b.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodertinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Harris B.R.;
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=9069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C. elegans: A platform for
investigating biology."
RL Science 282:2012-2018(1998).
DR EMBL; AL032657; CAD27614.1;
SQ SEQUENCE 1070 AA; 114180 MW; 75254D0DD5643AE5 CRC64;

Query Match 70.3%; Score 45; DB 5; Length 1070;
Best Local Similarity 70.0%; Pred. No. 4.9;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 CCDEYLOTFKE 11
Db 105 CCDEYLOTFKD 114

RESULT 3

Q9XWD6 PRELIMINARY; PRT; 1111 AA.
ID 09XWD6;
AC Q9XWD6;
DT 01-NOV-1999 (TReMBLrel. 12, Created)
DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Y47H9C.4 protein (CED-1).
GN Y47H9C.4 OR CED-1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodertinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Harris B.;
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Almscough R., Anderson K., Baynes C., Berks M.,
Bontfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
Craxton M., Dear S., Du Z., Durbin R., Favell A., Fulton L.,
Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
Parsons J., Percy C., Rifken L., Koopra A., Saunders D., Showkhen R.,
Smailton N., Smith A., Sonhammer E., Staden A., Sulston J.,
Therby-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterson R.,
Watson A., Weisstock L., Wilkinson-Sproat J., Wohlman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
elegans."
RL Nature 368:32-38(1994).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=21097720; PubMed=11163239;
RA Zhou Z., Hartwig E., Horvitz H.R.;
RT "CED-1 is a Transmembrane Receptor that Mediates Cell Corpse
Engulfment in C. elegans."
RL Cell 104:43-56(2001).

DR EMBL; AL032657; CAA21739.1; -;
DR EMBL; AF332568; AAG60061.1; -;
DR HSSP; P05106; 1Jv2.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR002174; Furin-like.
DR InterPro; IPR002049; Laminin_EGF.
DR Pfam; PF00008; EGF; 9.
DR PRINTS; PR00011; EGF_LAMININ.
DR SMART; SM00180; EGF_Lam; 6.
DR SMART; SM00001; EGF_Like; 5.
DR SMART; SM00261; FU; 2.
DR PROSITE; PS00022; EGF_1; UNKNOWN_15.
DR PROSITE; PS0186; EGF_2; 11.
KW EGF-like domain; Glycoprotein.
SQ SEQUENCE 1111 AA; 118803 MW; A39F374C008F9874 CRC64;

Query Match 70.3%; Score 45; DB 5; Length 1111;
Best Local Similarity 70.0%; Pred. No. 5.1;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 CCDEYLOTFKE 11
Db 105 CCDEYLOTFKD 114

RESULT 4

Q9BMO4 PRELIMINARY; PRT; 2304 AA.
ID 09BMO4;
AC 09BMO4;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE Putative BSC1 sodium channel protein.
OS Blattella germanica (German cockroach).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
OC Blaberoidea; Blattellidae; Blattellinae; Blattella.
OX NCBI_TaxID=6973;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21167530; PubMed=11267908;
RA Liu Z., Chung I., Dong K.;
RT "Alternative splicing of the BSC1 gene generates tissue-specific
RT isoforms in the German cockroach."
RL Insect Biochem. Mol. Biol. 31:703-713(2001).
DR EMBL; AF312365; AAK01090.1; -;
DR InterPro; IPR001682; Ca/Na_pore.
DR InterPro; IPR002111; Cat_channel_TrpL.
DR InterPro; IPR001064; Crystallin.
DR InterPro; IPR000566; Lipoch_cyFFABP.
DR InterPro; IPR000636; M+channel_nlg.
DR InterPro; IPR001696; Na_channel.
DR Pfam; PF00520; Ion_trans; 4.
DR PRINTS; PR00170; NACHANNEL.
DR PROSITE; PS00225; CRYSTALLIN_BETAGAMMA; UNKNOWN_1.
DR PROSITE; PS00213; LIPOCALIN; UNKNOWN_1.
KW Ionic channel.
SQ SEQUENCE 2304 AA; 260839 MW; 07D2D99AAAA611A45 CRC64;

Query Match 67.2%; Score 43; DB 5; Length 2304;
Best Local Similarity 75.0%; Pred. No. 24;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 CCDEYLOTF 9
Db 1264 CCDEYLOTF 1271

RESULT 5
Q56304 PRELIMINARY; PRT; 603 AA.
ID 056304;
AC 056304;
DT 01-JUN-1998 (TReMBLrel. 06, Created)

DT 01-JUN-1998 (TREMBlrel. 06, last sequence update)
 DE 01-DEC-2001 (TREMBlrel. 19, last annotation update)
 GN Envelope glycoprotein gp105.
 OS Human herpesvirus 7.
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Betaherpesvirinae.
 OX NCBI_TaxID=10372;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-RK:
 RX MEDLINE-98240941: PubMed=9581785;
 RA Megaw A.G., Kapaport D., Avidor B., Frenkel N., Davison A.J.;
 RT "The DNA sequence of the RK strain of human herpesvirus 7.";
 RL Virology 244:119-132(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-RK:
 RX Megaw A.G., Frenkel N., Davison A.J.;
 RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF037218; AAC40801.1; -
 SQ SEQUENCE 603 AA; 69500 MW; 37548A4161AD133F CRC64;

Query Match 65.6%; Score 42; DB 12; Length 603;
 Best Local Similarity 75.0%; Pred. No. 11;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 ACCDEYLQ 8
 DB 457 SCCDEYLK 464

RESULT 6
 ID 049696 PRELIMINARY; PRT; 560 AA.
 AC 049696:
 DT 01-JUN-1998 (TREMBlrel. 06, Created)
 DT 01-JUN-1998 (TREMBlrel. 06, last sequence update)
 DE 01-DEC-2001 (TREMBlrel. 19, last annotation update)
 GN Hypothetical 62.3 kDa protein.
 DT T6K21.150 OR AT4G17970
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Bevan M., Rieger M., Mueller-Auer S., Zipp M., Schaefer M.,
 RA Bancroft I., Mewes H.W., Mayer K., Schueller C.;
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Rieger M., Mueller-Auer S., Zipp M., Schaefer M., Mewes H.W.,
 RA Lemcke K., Mayer K.F.X.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL021889; CA117140.1; -
 DR EMBL; AL161547; CAB78799.1; -
 KW Hypothetical protein.
 SO SEQUENCE 560 AA; 62310 MW; 84587DFEB5100A8 CRC64;

Query Match 64.1%; Score 41; DB 10; Length 560;
 Best Local Similarity 63.6%; Pred. No. 15;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 ACCDEYLQTFKE 11
 DB 232 ACVDEYFEKE 242

RESULT 7
 ID 085MD0 PRELIMINARY; PRT; 742 AA.
 AC 085MD0:
 DT 01-JUN-2002 (TREMBlrel. 21, Created)
 DT 01-JUN-2002 (TREMBlrel. 21, last sequence update)
 DE 01-JUN-2002 (TREMBlrel. 21, last annotation update)
 GN Hypothetical protein ECU02_0870.
 DT ECU02_0870.
 OS Enecephalitozoon cuniculi.
 OC Eukaryota; Microsporidia; Unikaryonidae; Enecephalitozoon.
 OX NCBI_TaxID=6035;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-GB-M1;
 RA Genoscope;
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-GB-M1;
 RX MEDLINE-21576510: PubMed=11719806;
 RA Katinka M.D., Duprat S., Cornillot E., Metenier G., Thomarar F.,
 RA Prensier G., Barbe V., Peyretailade E., Brothier P., Wincker P.,
 RA Delbac F., El Alaoui H., Peyret P., Saurin W., Gouy M.,
 RA Weissbach J., Vivares C.P.;
 RT "Genome sequence and gene compaction of the eukaryote parasite
 RT Enecephalitozoon cuniculi.";
 RL Nature 414:450-453(2001).
 DR EMBL; AL590442; CAD25116.1; -
 SQ SEQUENCE 742 AA; 84381 MW; 2E2C38FE436E1321 CRC64;

Query Match 62.5%; Score 40; DB 5; Length 742;
 Best Local Similarity 100.0%; Pred. No. 30;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ACCDEY 6
 DB 155 ACCDEY 160

RESULT 8
 ID 095VR4 PRELIMINARY; PRT; 767 AA.
 AC 095VR4:
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, last sequence update)
 DE 01-MAR-2002 (TREMBlrel. 20, last annotation update)
 GN Hypothetical 89.6 kDa protein.
 DT F17N18.20 OR AT4G13130.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Bevan M., Pohl T., Weizenegger T., Bancroft I., Mewes H.W.,
 RA Mayer K.F.X., Mannhaupt G., Schueller C.;
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Pohl T., Weizenegger T., Mewes H.W., Lemcke K., Mayer K.F.X.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL: AL049751; CAB41923.1; -
DR EMBL: AL161535; CAB78355.1; -
DR InterPro: IPR004146; DC1.
DR InterPro: IPR000822; Znf_C2H2.
DR InterPro: IPR001965; Znf_PHD.
DR Pfam: PF03107; DC1; 3
DR SMART: SM00249; PHD; 3
DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
DR Hypothetical protein.
SQ SEQUENCE 767 AA; 89625 MW; 37A2D510FE1A8A9F CRC64;

Query Match 62.5%; Score 40; DB 10; Length 767;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CCDEYL 7
DB 280 CCDEYL 285
|||||

RESULT 9
Q9ZCM9 PRELIMINARY; PRT; 227 AA.
AC Q9ZCM9;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Hypothetical protein RP694.
GN RP694.
OS Rickettsia prowazekii.
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC Rickettsiaceae; Rickettsiae; Rickettsia.
OX NCBI_TaxID=782;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MADRID E;
RX MEDLINE=9903949; PubMed=9823893;
RA Andersson S.G.E., Zomorodipour A., Andersson J.O.,
RA Scheritz-Ponten T., Alsmark U.C.M., Podowski R.M., Naeslund A.K.,
RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
RT "The genome sequence of Rickettsia prowazekii and the origin of
RT mitochondria."
RL Nature 396:133-140(1998).
DR EMBL: AJ235272; CA15131.1; -
DR InterPro: IPR004479; ExsB.
DR TIGRPFAM: TIGR00364; ExsB; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 227 AA; 25211 MW; 7EB608D1261BE081 CRC64;

Query Match 60.9%; Score 39; DB 16; Length 227;
Best Local Similarity 60.0%; Pred. No. 15;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 CCDEYLQTK 11
DB 136 CCDEYLQSF 145
|||||

RESULT 10
Q8RYT0 PRELIMINARY; PRT; 503 AA.
AC Q8RYT0;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Putative ubiquitin-specific protein.
GN B156H12.20.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliopsida; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN=CV, NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, BAC
RT clone:B1156H12."
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF004225; BAB90685.1; -
SQ SEQUENCE 503 AA; 56469 MW; E7A941D7305366EC CRC64;

Query Match 60.9%; Score 39; DB 10; Length 503;
Best Local Similarity 66.7%; Pred. No. 33;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ACCDEYLQ 9
DB 418 ACCDPYFNT 426
|||||

RESULT 11
Q9U5F5 PRELIMINARY; PRT; 537 AA.
AC Q9U5F5;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE RNA polymerase III largest subunit (Fragment).
OS Eptatretus burgeri (Inshore hagfish).
OC Eukaryota; Metazoa; Chordata; Hyperotreti; Myxiniiformes;
OC Myxiniidae; Eptatretinae; Eptatretus.
OX NCBI_TaxID=7764;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE=20063780; PubMed=10594174;
RA Kuraku S., Hoshiyama D., Katoh K., Suga H., Miyata T.;
RT "Wonyhyly of lampreys and hagfishes supported by nuclear DNA-coded
RT genes."
RL J. Mol. Evol. 49:729-735(1999).
DR EMBL: AB025333; BAA8486.1; -
DR InterPro: IPR000722; RNA_pol_A.
DR InterPro: IPR002879; RNA_pol_A2.
DR Pfam: PF00623; RNA_pol_A; 1.
DR Pfam: PF01854; RNA_pol_A2; 1.
FT NON_TER 1 537
FT TER 537
SQ SEQUENCE 537 AA; 58125 MW; 24F34D89483DE900 CRC64;

Query Match 60.9%; Score 39; DB 13; Length 537;
Best Local Similarity 55.6%; Pred. No. 35;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 CCDEYLQTK 10
DB 513 CCDEYLQAR 521
|||||

RESULT 12
Q9F1Z5 PRELIMINARY; PRT; 87 AA.
AC Q9F1Z5;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE BtrI.
GN BTRI.
OS Bacillus circulans.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Bacillaceae; Bacillus.
OX NCBI_TaxID=1397;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21019014; PubMed=11132962;
RA Ota Y., Tamegai H., Kudo F., Kuriki H., Kolke-Takeshita A., Eguchi T.,

```
RA kakiuma K.;
RT "Butirosin-biosynthetic Gene Cluster from Bacillus circulans.";
RL J. Antibiot. 53:1158-1167(2000).
DR EMBL: AB033991; BAB18048.1; -.
DR InterPro: IPR003880; Praprine_attach.
DR Pfam: PF00550; pp-binding. 1.
DR PROSITE: PS00075; ACP_DOMAIN. 1.
DR PROSITE: PS00012; PHOSPHOPANTETHEINE; UNKNOWN_1.
KW Phosphopantetheine.
SQ SEQUENCE 87 AA; 9781 MW; FA15D023B2003DC8 CRC64;

Query Match
Best Local Similarity 59.4%; Score 38; DB 2; Length 87;
Pred. No. 9.7;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACCEYLQTRKE 11
DB 16 SCVNEILQTRKE 26

RESULT 13
ID 09VF42 PRELIMINARY; PRT; 219 AA.
AC 09VF42;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE CG5623 protein.
GN CG5623.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OC NCBI_TaxId=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutcon G.G., Mortan J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abiri J.F., Agbayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibeagwa C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laslo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mlshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleib J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spindling A.C., Stapleton M., Strong R., Sun E.,
RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
```

```
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL: AE003710; AAF55218.1; -.
DR FlyBase: FBgn0038357; CG5623.
SQ SEQUENCE 219 AA; 25031 MW; 68240DEA42857440 CRC64;

Query Match
Best Local Similarity 59.4%; Score 38; DB 5; Length 219;
Pred. No. 23;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCEYLQTR 9
DB 211 ACCEDLQDS 219

RESULT 14
ID 08T485 PRELIMINARY; PRT; 279 AA.
AC 08T485;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE A010953p.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OC NCBI_TaxId=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guartin H., Kronmiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleib J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY089306; AAL90044.1; -.
SQ SEQUENCE 279 AA; 31990 MW; 5AF234E59ACEF0B4 CRC64;

Query Match
Best Local Similarity 59.4%; Score 38; DB 5; Length 279;
Pred. No. 29;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCEYLQTR 9
DB 271 ACCEDLQDS 279

RESULT 15
ID P79952 PRELIMINARY; PRT; 340 AA.
AC P79952;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE xfringe2.
GN XFRINGE2.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodidae; Xenopus.
OX NCBI_TaxId=8355;
RN [1]
RP SEQUENCE FROM N.A.
RA Fan M.J., Sokol S.Y.;
RA Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RL EMBL: U9750; AAB38363.1; -.
DR InterPro: IPR003378; Frlng.
DR InterPro: IPR000822; Znf_C2H2.
DR Pfam: PF02434; Frlng. 1.
DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
```

SO SEQUENCE 340 AA; 38570 MM; 882679543416B5D7 CRC64;

Query Match 59.4%; Score 38; DB 13; Length 340;

Best Local Similarity 66.7%; Pred. No. 35;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 CCDEYLOTK 10

11:11:11

Db 45 CSDKYLETK 53

Search completed: November 18, 2002, 15:58:10
Job time : 27.1935 secs

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV, COLUMBIA;
 RX MEDLINE=21016719; PubMed=1130712;
 RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
 RA White O., Alonso J., Altati H., Araujo R., Bowman C.L., Brooks S.Y.,
 RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
 RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
 RA Dunn P., Egu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
 RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Hultzer L.,
 RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
 RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
 RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
 RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marzilli A.,
 RA Miltischer J., Miranda M., Nguyen M., Niernm W.C., Osborne B.I.,
 RA Pal G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
 RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
 RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
 RA Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
 RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.,
 RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
 thaliana."
 RT Nature 408:816-820(2000).
 DR EMBL: AC037424; AM51543.1;
 SQ SEQUENCE 784 AA; 87099 MW; B05412253B09C76F CRC64;

Query Match 76.7%; Score 33; DB 10; Length 784;
 Best Local Similarity 66.7%; Pred. No. 2.9e+02;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 NMSEVERE 9
 Db 438 SMLEIERE 446
 :|||:||||

RESULT 14
 ID 080266 PRELIMINARY; PRT; 59 AA.
 AC 080266;
 DT 01-JUN-2002 (TREMBLrel. 21, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Hypothetical protein PF0984.
 GN PF0984.
 OS Pyrococcus furiosus.
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
 OC Pyrococcus.
 OX NCBI_TaxID=2261;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=VCI / DSM 3638 / ATCC 43587 / JCM 8422;
 RA Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.,
 RT "The complete sequence of the Pyrococcus furiosus genome."
 RL Submitted (FEF-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF010210; AAL81108.1;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 59 AA; 7013 MW; EA5EBF0E27533CA CRC64;

Query Match 74.4%; Score 32; DB 17; Length 59;
 Best Local Similarity 66.7%; Pred. No. 39;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0.

OY 1 NMSEVERE 9
 Db 49 NMLKEVERD 57
 :|||:||||

RESULT 15
 09V101

ID 09V101 PRELIMINARY; PRT; 222 AA.
 AC 09V101;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE CG10659 protein.
 GN CG10659.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abrell J.F., Agbayani A., An H.-J., Andrews-Plantkock C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhattacharya D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotier P.,
 RA Burks K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fouts K., Gabrielian A.E., Gary N.S., Galardi W.M., Glasser K.,
 RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,
 RA Lascko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Misha N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RT "The genome sequence of Drosophila melanogaster."
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003665; AAF53864.1;
 DR Flybase: FBgn0032839; CG10659.
 SQ SEQUENCE 222 AA; 24902 MW; 22A5D6114BFD67A CRC64;

Query Match 74.4%; Score 32; DB 5; Length 222;
 Best Local Similarity 75.0%; Pred. No. 1.4e+02;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0.

OY 2 MUSEVERE 9
 Db 112 MLEKIERE 119
 :|||:||||

Search completed: November 18, 2002, 15:58:08
 Job time : 22.6129 secs

```

DR Interpro: IPR003593; AAA_ATPase.
DR Pfam: PF00005; ABC_tran; 1.
DR ProDom: PD000006; ABC_transportr; 1.
DR SMART: SM00382; AAA; 1.
DR PROSITE: PS00211; ABC_TRANSPORTER; 1.
KM ATP-binding; Transport.
FT NON_TER 284
SQ SEQUENCE 284 AA; 31759 MW; 30AE30A9C6AD3F90 CRC64;

Query Match
Best Local Similarity 76.7%; Score 33; DB 2; Length 284;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NMLSEVERE 9
Db 208 NMLSEVERE 216

RESULT 10
QY1185 PRELIMINARY; PRT; 284 AA.
AC QY1185;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DE Hypothetical protein PA2395.
GN PA2395.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=20437337; Pubmed=10984043;
RX STRAIN=ATCC 15692 / PA01;
RA Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.L., Huynh W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen.";
RL Nature 406:959-964(2000).
DR EMBL: AE004665; AAC05783.1; -.
DR InterPro: IPR001412; TRNA-synt_1.
DR PROSITE: PS00178; AA_TRNA_LIGASE_I; UNKNOWN_1.
KM Hypothetical protein: Complete proteome.
SQ SEQUENCE 284 AA; 31070 MW; 1916488599EA777 CRC64;

Query Match
Best Local Similarity 76.7%; Score 33; DB 16; Length 284;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 NMLSEVERE 9
Db 149 NMLSEVERE 156

RESULT 11
QY7LF8 PRELIMINARY; PRT; 721 AA.
AC QY7LF8;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Superfamily I DNA and RNA helicase.
GN CAC0603.
OS Clostridium acetobutylicum.
OC Bacteria; Firmicutes; Bacilli/Clostridium group; Clostridia;
OC Clostridiales; Clostridiaceae; Clostridium.
OX NCBI_TaxID=1488;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX MEDLINE=21359325; Pubmed=11466286;
RA Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
RA Gibson R., Lee H.M., Dubois J., Qiu D., Hltl J., Wolf Y.I.,
RA Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
RA Bennett G.N., Koonin E.V., Smith D.R.;
RT "Genome sequence and comparative analysis of the solvent-producing
RT bacterium Clostridium acetobutylicum.";
RL J. Bacteriol. 183:4823-4838(2001).
DR EMBL: AE007575; AAK78581.1; -.
DR InterPro: IPR000212; UvrD-helicase.
DR Pfam: PF00580; UvrD-helicase; 2.
KM Helicase; Complete proteome.
SQ SEQUENCE 721 AA; 84757 MW; FA9B2B134B1839PF CRC64;

Query Match
Best Local Similarity 76.7%; Score 33; DB 16; Length 721;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 NMLSEVERE 9
Db 349 NMLSEVERE 357

RESULT 12
QY4AE1 PRELIMINARY; PRT; 755 AA.
AC QY4AE1;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE At1g52410/F19K6.14.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC Eucosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Shih P., Chen H., Cheuk R., Kim C.J., Koesema E., Meyers M.C.,
RA Banb J., Bowser L., Carninci P., Dale J.M., Gibson H.A.,
RA Goldsmith A.D., Hayashizaki Y., Ishida J., Jiang P.X., Jones T.,
RA Kamiya A., Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J.,
RA Liu S.X., Miranda M., Natsusaka M., Nguyen M., Onodera C.S., Palm C.J.,
RA Pam P.K., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
RA Tang C.C., Toriumi M., Yamada K., Yu G., Yu S., Shinozaki K.,
RA Davis R.M., Theologis A., Ecker J.R.;
RT "Arabidopsis cDNA clones.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY048247; AAK82509.1; -.
SQ SEQUENCE 755 AA; 84046 MW; 740F880DEB1AE2D2 CRC64;

Query Match
Best Local Similarity 76.7%; Score 33; DB 10; Length 755;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NMLSEVERE 9
Db 409 NMLSEVERE 417

RESULT 13
QYCB31 PRELIMINARY; PRT; 784 AA.
AC QYCB31;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE Myosin-like protein.
GN F19K6.14.
OS Arabidopsis thaliana (Mouse-ear cress).

```


DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Tumor necrosis factor.
 OS Paratichthys olivaceus (Flounder).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;
 OC Pleuronectidae; Paratichthyidae; Paratichthys.
 NCBI_TaxID=8255;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20491932; PubMed=11035080;
 RA Hiroo I., Nam B., Kurobe T., Aoki T.;
 RT "Molecular Cloning, Characterization, and Expression of TNF cDNA and
 RT Gene from Japanese Flounder Paratichthys olivaceus.";
 RL J. Immunol. 165:4423-4427(2000).
 DR EMBL; AB040449; BAA94970.1; -.
 DR HSSP; P01375; 4TSV.
 DR InterPro: IPR003636; TNF_family.
 DR InterPro: IPR000478; TNF_family.
 DR Pfam: PF00229; TNF_1.
 DR PRINTS: PR01234; TNECROSISFCT.
 DR ProDom: PD002012; TNF_abc; 1.
 DR SMART; SM00207; TNF; 1.
 DR PROSITE; PS50049; TNF_2; 1.
 SQ SEQUENCE 225 AA; 24889 MW; 7E347FA253BF3643 CRC64;

Query Match 79.1%; Score 34; DB 13; Length 225;
 Best Local Similarity 77.8%; Pred. No. 55;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 NMUSEVERE 9
 Db 206 NMUSELETE 214

RESULT 7
 ID 09RSZ3 PRELIMINARY; PRT; 766 AA.
 AC 09RSZ3;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-JUN-2000 (TREMBlrel. 21, Last annotation update)
 DE DNA mismatch repair protein MMRs, putative.
 GN DRI1976.
 OS Delnoccocus radiodurans;
 OC Bacteria; Thermus/Delnoccocus group; Delnoccoci; Delnoccocales;
 OC Delnoccocaceae; Delnoccoccus.
 NCBI_TaxID=1299;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=RI;
 RX MEDLINE=20036896; PubMed=10567266;
 RA White O., Eissen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
 RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
 RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
 RA Vamathevan J.D., Lam P., McDonald L., Uitterback T., Zaleski C.,
 RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
 RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
 RA Fraser C.M.;
 RT "Genome sequence of the radioresistant bacterium Delnoccoccus
 RT radiodurans RI.";
 RL Science 286:1571-1577(1999).
 DR EMBL; AE002035; AAF11527.1; -.
 DR TIGR; DRI1976; -.
 DR InterPro: IPR000432; Muts_C.
 DR InterPro: IPR002863; Muts_N.
 DR InterPro: IPR002625; Smr/Muts2.
 DR Pfam: PF00488; Muts_C; 1.
 DR Pfam: PF01713; Smr; 1.
 DR ProDom: PD001263; Muts_C; 1.
 DR SMART; SM00534; Mutsac; 1.

DR SMART; SM00533; Mutsd; 1.
 DR SMART; SM00463; SMR; 1.
 DR TIGRFAMs; TIGR01069; muts2; 1.
 KW Complete proteome.
 SQ SEQUENCE 766 AA; 84141 MW; 361E231F4B47C41D CRC64;

Query Match 79.1%; Score 34; DB 16; Length 766;
 Best Local Similarity 87.5%; Pred. No. 18e+02;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 MUSEVERE 9
 Db 513 MUSELERE 520

RESULT 8
 ID 09BNR1 PRELIMINARY; PRT; 124 AA.
 AC 09BNR1;
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DE RNA polymerase II largest subunit (Fragment).
 OS Chaetopteleura apiculata (Common eastern chiton).
 OC Eukaryota; Metazoa; Mollusca; Polyplacophora; Neoloricata;
 OC Ischnochitonina; Chaetopteleuridae; Chaetopteleura.
 NCBI_TaxID=58794;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Regier J.C., Shultz J.W.;
 RT "Resolving arthropod relationships with elongation factor-2,
 RT elongation factor-1alpha, and RNA polymerase II sequences.";
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBD databases.
 DR EMBL; AF240884; AKK12363.1; -.
 DR InterPro: IPR000722; RNA_pol_A.
 DR Pfam: PF00623; RNA_pol_A; 1.
 FT NON_TER 1
 FT TER 124
 SQ SEQUENCE 124 AA; 14456 MW; 37A4C25A76A56D5 CRC64;

Query Match 76.7%; Score 33; DB 5; Length 124;
 Best Local Similarity 66.7%; Pred. No. 50;
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 NMUSEVERE 9
 Db 86 NVLADVERE 94

RESULT 9
 ID P72410 PRELIMINARY; PRT; 284 AA.
 AC P72410;
 DT 01-FEB-1997 (TREMBlrel. 02, Created)
 DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Putative ABC transporter intracellular ATPase subunit BldKE
 DE (Fragment).
 GN BldKE.
 OS Streptomyces coelicolor.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomycetes.
 NCBI_TaxID=1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=M145;
 RA Nowell J.R., McGovern K., Losick R.;
 RT "An oligopeptide permease responsible for the import of an
 RT extracellular signal governing aerial mycelium formation in
 RT Streptomyces coelicolor.";
 RL Submitted (AUG-1996) to the EMBL/GenBank/DBD databases.
 CC -i- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
 DR EMBL; U68036; AAB09558.1; -.

FT VARSPLIC 1856 1869 MISSING (IN ISOFORM 4 AND ISOFORM 5).
 FT CONFLICT 294 297 CLRO -> SMO (IN REF. 1).
 FT CONFLICT 2262 2262 L -> I (IN REF. 1).
 FT CONFLICT 2530 2530 I -> S (IN REF. 1).
 FT CONFLICT 2907 2907 A -> P (IN REF. 1).
 FT CONFLICT 3194 3194 L -> Q (IN REF. 1).
 FT CONFLICT 3341 3341 A -> G (IN REF. 1).
 FT CONFLICT 4076 4076 T -> I (IN REF. 1).
 SQ SEQUENCE 5107 AA; 578839 MW; 2AC48FDC058F36F CRC64;
 Query Match 88.4%; Score 38; DB 5; Length 5107;
 Best Local Similarity 88.9%; Pred. No. 17e+02;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 NMLSEVERE 9
 Db 2871 NMLSEVERE 2879
 RESULT 3
 Q96V77 PRELIMINARY; PRT; 724 AA.
 AC 096V77;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Phenylalanine ammonia-lyase (EC 4.3.1.5).
 GN PAL1.
 OS Ustilago maydis (Smut fungus).
 OC Eukaryota; Fungi; Basidiomycota; Ustilaginomycetes;
 OC Ustilaginomycetidae; Ustilaginales; Ustilaginaceae; Ustilago.
 OX NCBI_TaxID=5270;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kim S.-H., Macdonald K., Virmani D., Wake K., Kronstad J.W.,
 RA Ellis B.E.;
 RT "Cloning and disruption of a phenylalanine ammonia-lyase gene from
 RT Ustilago maydis.";
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF306551; AAL09388.1; -.
 DR InterPro; IPR003006; IG_MHC.
 DR InterPro; IPR001106; Phe/His_NH3lyase.
 DR Pfam; PF00221; PAL; 1.
 DR TIGRFAMs; TIGR01226; phe_am_lyase; 1.
 DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
 KW Lyase.
 SQ SEQUENCE 724 AA; 79277 MW; 02A73C97C34CD1AB CRC64;
 Query Match 86.0%; Score 37; DB 3; Length 724;
 Best Local Similarity 77.8%; Pred. No. 41;
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NMLSEVERE 9
 Db 715 NMLSEVERE 723
 RESULT 4
 Q8T100 PRELIMINARY; PRT; 209 AA.
 AC 08T100;
 DT 01-JUN-2002 (TREMBlrel. 21, Created)
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Predicted protein.
 GN MA4358.
 OS Methanosarcina acetivorans.
 OC Archaea; Euryarchaeota; Methanococci; Methanosarcinales;
 OC Methanosarcinaceae; Methanosarcina.
 OX NCBI_TaxID=2214;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C2A / ATCC 35395 / DSM 2834;

RX MEDLINE-21929760; PubMed-11932238;
 RA Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., Macdonald P.,
 RA Fitzhugh W., Calvo S., Engels R., Smirnov S., Anoor D., Brown A.,
 RA Allen N., Naylor J., Stange-Thomann N., Dearellano K., Johnson R.,
 RA Jinton L., McKernan P., McKernan K., Talamas J., Tirrell A., Ye W.,
 RA Zimmer A., Barber R.D., Cann I., Graham D.E., Grahame D.A., Guss A.M.,
 RA Hedderich R., Ingram-Smith C., Kueltner H.C., Krzycki J.A.,
 RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
 RA Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,
 RA Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,
 RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
 RA Metcalf W.W., Birren B.;
 RT "The genome of Methanosarcina acetivorans reveals extensive metabolic
 RT and physiological diversity.";
 RL Genome Res. 12:532-542(2002).
 DR EMBL; AE011154; AAM07700.1; -.
 KW Complete proteome.
 SQ SEQUENCE 209 AA; 23133 MW; C664EDC256615C1B CRC64;
 Query Match 79.1%; Score 34; DB 17; Length 209;
 Best Local Similarity 66.7%; Pred. No. 51;
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NMLSEVERE 9
 Db 60 NMLSEVERE 68
 RESULT 5
 Q91B42 PRELIMINARY; PRT; 225 AA.
 AC 091B42;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Tumor necrosis factor.
 OS Paralicthys olivaceus (Flounder).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;
 OC Pleuronectoidae; Paralicthyidae; Paralicthys.
 OX NCBI_TaxID=8255;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE-20491932; PubMed-11035080;
 RA Hirono I., Nam B., Kurobe T., Aoki T.;
 RT "Molecular Cloning, Characterization, and Expression of TNF cDNA and
 RT Gene from Japanese Flounder Paralicthys olivaceus.";
 RL J. Immunol. 165:4423-4427(2000).
 DR EMBL; AB040448; BAA94969.1; -.
 DR HSSP; P01375; 4TSV.
 DR InterPro; IPR003636; TNF_abc.
 DR InterPro; IPR000478; TNF_family.
 DR Pfam; PF00229; TNF; 1.
 DR PRINTS; PR01234; TNECROSISFCT.
 DR ProDom; PD002012; TNF_abc; 1.
 DR SMART; SM00207; TNF; 1.
 DR PROSITE; PS50049; TNF_2; 1.
 SQ SEQUENCE 225 AA; 24965 MW; 8F947FB25FC82658 CRC64;
 Query Match 79.1%; Score 34; DB 13; Length 225;
 Best Local Similarity 77.8%; Pred. No. 55;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 NMLSEVERE 9
 Db 206 NMLSEVERE 214
 RESULT 6
 Q91B41 PRELIMINARY; PRT; 225 AA.
 AC 091B41;
 Q91B41;

ID	Q24498	PRELIMINARY	PRT: 5107 AA.
AC	Q24498; Q24499; Q24500; Q24501; Q9VAY7;		
DT	01-NOV-1996 (TREMBLrel. 01, Created)		
DR	01-JUN-2000 (TREMBLrel. 14, Last sequence update)		
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)		
DE	Ryanodine receptor.		
GN	RYA-R44F OR CG10844.		
OS	Drosophila melanogaster (Fruit fly).		
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;		
CC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;		
CC	Ephydroidea; Drosophilidae; Drosophila.		
OK	NCBI_TaxID=72277;		
RN	[1]		
RX	SEQUENCE FROM N.A. (ISOFORMS 2; 3; 4 AND 5), AND CHARACTERIZATION.		
RP	MEDLINE=94102409; Pubmed=8276118;		
RA	Takeshima H., Nishi M., Iwabe N., Miyata T., Hosoya T., Masai I.,		
RT	Hotta Y.		
RL	"Isolation and characterization of a gene for a ryanodine		
RL	receptor/calcium release channel in Drosophila melanogaster.";		
RL	FEBS Lett. 337:81-87(1994).		
RN	[2]		
RP	SEQUENCE FROM N.A. (ISOFORM 1).		
RC	STRAIN-BERKELEY.		
RX	MEDLINE=20196006; Pubmed=10731132;		
RA	Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,		
RA	Amantides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,		
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,		
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,		
RA	Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,		
RA	Wan K.H., Doyle C., Baxter E.G., Helt J., Nelson C.R., Miklos G.L.G.,		
RA	Abdil J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,		
RA	Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,		
RA	Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,		
RA	Borkov D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,		
RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,		
RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,		
RA	de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,		
RA	Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,		
RA	Dubin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,		
RA	Foster C., Gabrielian A.E., Gang N.S., Gelbart W.M., Glasser K.,		
RA	Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,		
RA	Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,		
RA	Hosltin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,		
RA	Jastal M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,		
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,		
RA	Laoko P., Lei Y., Levitsky A.C., Li J., Li Z., Liang Y., Lin X.,		
RA	Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,		
RA	Merkulov G., Mishina N.V., Mobarry A.C., Morris J., Moshrefi A.,		
RA	Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,		
RA	Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Paclab J.M.,		
RA	Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,		
RA	Reinett K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,		
RA	Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,		
RA	Spieler E., Spradling A.C., Stapleton M., Strong R., Sun E.,		
RA	Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,		
RA	Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,		
RA	Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,		
RA	Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhu S., Zhu Q., Zheng L.,		
RA	Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Zheng L.,		
RA	Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;		
RT	"The genome sequence of Drosophila melanogaster.";		
RL	Science 287:2185-2195(2000).		
CC	-1- FUNCTION: COMMUNICATION BETWEEN TRANSVERSE-TUBULES AND		
CC	SARCOPLASMIC RETICULUM. CONTRACTION OF MUSCLE IS TRIGGERED BY		
CC	RELEASE OF CA++ FROM SR FOLLOWING DEPOLARIZATION OF T-TUBULES (BY		
CC	STIMULATY).		
CC	-1- FUNCTION: INTRACELLULAR CALCIUM CHANNEL. MAY BE INVOLVED IN		
CC	MORPHOGENETIC MOVEMENTS DURING EMBRYONIC DEVELOPMENT.		
CC	-1- SUBUNIT: HOMOTETRAMER (POTENTIAL).		
CC	-1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).		
CC	-1- ALTERNATIVE PRODUCTS: 5 ISOFORMS; 1 (SHOWN HERE), 2, 3, 4 AND 5;		

[illegible]

VTKMHAGSIEHFVRYEAGDVIGCFIDVKEQ -> EEKVYGG
GVSESEFGKQCGPDIIVGVFLDLADH (IN ISOFORM 3
AND ISOFORM 4).

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 18, 2002, 15:53:18 ; Search time 20.6129 Seconds
(without alignments)
89.964 Million cell updates/sec

Title: US-09-848-781-1
Perfect score: 43
Sequence: 1 NMLSEVERE 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPREMBL.21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertbrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	38	88.4	772	10	Q9LSB4
2	38	88.4	5107	5	Q24498
3	37	86.0	724	3	Q96V77
4	34	79.1	209	17	Q8RT00
5	34	79.1	225	13	Q9IB42
6	34	79.1	225	13	Q9IB41
7	34	76.7	766	16	Q9RSZ3
8	33	76.7	124	5	Q9BNR1
9	33	76.7	284	2	P72410
10	33	76.7	284	16	Q9I185
11	33	76.7	721	16	Q97LF8
12	33	76.7	755	10	Q94AE1
13	33	76.7	784	10	Q9C831
14	32	74.4	59	17	Q8U266
15	32	74.4	222	5	Q9V101
16	32	74.4	331	2	Q8VL29

17	32	74.4	388	10	Q9LZU0	Q9LZU0 arabidopsis
18	32	74.4	409	3	Q05543	Q05543 saccharomyc
19	32	74.4	428	16	Q8RHS8	Q8RHS8 fusobacteri
20	32	74.4	578	3	Q9Y7C4	Q9Y7C4 candida alb
21	32	74.4	636	5	Q45229	Q45229 caenorhabdi
22	32	74.4	645	5	Q45943	Q45943 caenorhabdi
23	32	74.4	1088	5	Q00905	Q00905 oxytricha f
24	32	74.4	1798	4	Q14997	Q14997 homo sapien
25	32	74.4	2556	12	Q90H56	Q90H56 gallid herp
26	32	74.4	2656	5	Q9NE92	Q9NE92 leishmania
27	31	72.1	114	9	Q8SCV0	Q8SCV0 pseudomonas
28	31	72.1	132	16	Q8RH51	Q8RH51 yersinia pe
29	31	72.1	143	16	Q93MC6	Q93MC6 clostridium
30	31	72.1	190	16	Q9ZCQ2	Q9ZCQ2 rickettsia
31	31	72.1	190	16	Q8ZJ75	Q8ZJ75 yersinia pe
32	31	72.1	200	10	Q9LXH5	Q9LXH5 arabidopsis
33	31	72.1	277	17	Q59137	Q59137 pyrococcus
34	31	72.1	283	5	Q9VQ14	Q9VQ14 drosophila
35	31	72.1	284	5	Q95SG5	Q95SG5 drosophila
36	31	72.1	314	16	Q8RCX7	Q8RCX7 thermoplas
37	31	72.1	326	17	Q97CS8	Q97CS8 thermoplas
38	31	72.1	333	16	Q9JZL5	Q9JZL5 neisseria m
39	31	72.1	341	5	Q9XU72	Q9XU72 caenorhabdi
40	31	72.1	353	17	Q28776	Q28776 archaeoglob
41	31	72.1	366	16	Q9Z7Y8	Q9Z7Y8 chlamydia p
42	31	72.1	426	3	Q9C1K7	Q9C1K7 neurospora
43	31	72.1	429	17	Q980V3	Q980V3 sulfobolus
44	31	72.1	482	10	Q81077	Q81077 arabidolum
45	31	72.1	540	5	Q96320	Q96320 plasmodium

ALIGNMENTS

RESULT 1

Q9LSB4 ID Q9LSB4 PRELIMINARY: PRT: 772 AA.

AC Q9LSB4; 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 15, Last annotation update)
DE Genomic DNA, chromosome 3, pl clone: MVC8.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RA Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Tabata S.;
RL Submitted (Apr-1999) to the EMBL/GenBank/DBJ databases.

RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RX MEDLINE=20277480; PubMed=10819329;
RA Nakamura Y.;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence
features of the regions of 4,504,864 bp covered by sixty pl and TAC
RT clones.";

RL DNA Res. 7.131-135(2000).
DR EMBL: AB026553; BAB02881.1; -
SQ SEQUENCE 772 AA; 85016 MW; 8A2CD740BDA557A6 CRC64;

Query Match 88.4%; Score 38; DB 10; Length 772;

Best Local Similarity 77.8%; Pred. No. 28;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NMLSEVERE 9
DB 263 NMLSEVERE 271

PA (EXPR-) EXPRESSIVE CONSTRUCTS INC.
XX Sanders MC;
XX
XX
DR WPI: 2002-303940/34.
XX
XX Detection of bacterial contamination in foods or food-related work
PT areas comprises identifying a protein specific to the microorganism
PT using fluorescence or colorimetric methods -
XX
XX
PS Disclosure; Page 7; 25pp; English.
XX
XX The invention describes a specific and sensitive method for detection of
CC a pathogenic microorganism in potentially contaminated food products
CC at the retail level, by detecting a bacteria-specific protein e.g. a
CC *Listeria monocytogenes* metalloprotease. The process is useful for
CC detecting *Listeria monocytogenes*, a common cause of gastroenteritis, or
CC other food contaminants in food products or food-related work areas.
CC This sequence represents the peptide M1 that functions as a substrate
CC for the *Listeria monocytogenes* metalloprotease mpl. The peptide is
CC specific to *L. monocytogenes* at pH 5.5 and can be used as a probe for
CC fluorescence resonance energy transfer (FRET) analysis of mpl activity
CC in food products.
XX
XX Sequence 9 AA:
SQ
Query Match 100.0%; Score 43; DB 23; Length 9;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 NMLSEVERE 9
DB 1 NMLSEVERE 9
RESULT 2
ABBA47678
ID ABBA47678 standard; Protein; 510 AA.
XX
XX ABBA47678;
AC
XX
XX 05-FEB-2002 (first entry)
DT
XX
XX *Listeria monocytogenes* protein #382.
DE
XX
XX Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
KM vitamin B12; bacterial infection; disease.
XX
XX *Listeria monocytogenes*.
OS
XX
XX WO200177335-A2.
PN
XX
XX 18-OCT-2001.
PD
XX
XX 11-APR-2001; 2001WO-FR01118.
PE
XX
XX 11-APR-2000; 2000FR-0004629.
PR
XX
XX (INSP) INST PASTEUR.
PA
XX Buchrieser C, Frangeul L, Couve E, Rusnlok C, Fahl H, Dehoux P,
PI Dussurgeat O, Chelouani F, Nedjati H, Glaser P, Kunst F, Cossart P;
PI Daniels J, Goebel W, Kreft J, Kuhn M, Ng E, Vazquez-Boland JA;
PI Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A;
PI Chakraborty T, Dommann E, Hain T, Berche P, Charbit A, Durant L;
PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N;
PI Madueno E, De Pablos B, Wehlund J, Kaerst U, Entian K, Hauf J;
PI Rose M, Voss H;
XX
XX WPI: 2002-010914/01.
DR
XX
XX Genomic sequence for *Listeria monocytogenes*, useful e.g. for treatment
PT and prevention of *Listeria* and related bacterial infections, and

PT related polypeptides -
XX
XX
PS Claim 6; SEQ ID NO 383; 192pp; French.
XX
XX
XX The present invention relates to the genome sequence of *Listeria*
CC *monocytogenes* EGD-e (see ABA03041). The genome sequence and fragments of
CC it are useful for selecting probes and primers for detecting genes in *L.*
CC *monocytogenes* and related organisms, and for studying genetic
CC polymorphisms and other genomes. The present sequence is a protein
CC encoded by the genome sequence of the present invention. Proteins
CC expressed from the genome sequence are useful for raising specific
CC antibodies, identification of *L. monocytogenes* and related organisms, and
CC for biosynthesis and biodegradation, especially biosynthesis of Vitamin
CC B12. The genome sequence and proteins encoded by it are also useful for
CC selecting compounds that regulate gene expression and cell replication
CC and modulate *L. monocytogenes*-related diseases. In addition, the genome
CC sequence and proteins encoded by it are useful in pharmaceutical and
CC vaccines compositions for the treatment or prevention of infections by *L.*
CC *monocytogenes* and related organisms.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 510 AA:
SQ
Query Match 88.4%; Score 38; DB 23; Length 510;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 NMLSEVER 8
DB 196 NMLSEVER 203
RESULT 3
ABBA65257
ID ABBA65257 standard; Protein; 5107 AA.
XX
XX ABBA65257;
AC
XX
XX 26-MAR-2002 (first entry)
DT
XX
XX *Drosophila melanogaster* polypeptide SEQ ID NO 22563.
DE
XX
XX *Drosophila*; developmental biology; cell signalling; insecticide;
KM pharmaceutical.
XX
XX *Drosophila melanogaster*.
OS
XX
XX WO200171042-A2.
PN
XX
XX 27-SEP-2001.
PD
XX
XX 23-MAR-2001; 2001WO-US09231.
PE
XX
XX 23-MAR-2000; 2000US-191637P.
PR
XX 11-JUL-2000; 2000US-0614150.
PR
XX
XX (PEKE) PE CORP NY.
PA
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
PI
XX
XX WPI: 2001-656860/75.
DR N-PSDB; ABL09360.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from *Drosophila* and for elucidating cell signalling and cell-cell
PT interactions -
XX
XX
XX Disclosure; SEQ ID NO 22563; 21pp + Sequence Listing; English.
PS
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from *Drosophila*. The invention is

CC	useful in developmental biology and in elucidating cell signalling and
CC	cell-cell interactions in higher eukaryotes for the development of
CC	insecticides, therapeutics and pharmaceutical drugs. The invention
CC	discloses genomic DNA sequences (AB116176-AB130511), expressed DNA
CC	sequences (AB101840-AB16175) and the encoded proteins
CC	(AAB57737-AB872072).
CC	The sequence data for this patent did not form part of the printed
CC	specification, but was obtained in electronic format directly from WIPO
CC	at ftp.wipo.int/pub/published_pct_sequences.
CC	
SO	Sequence 5107 AA:
Oy	1 NWLSEVERE 9
Db	2871 NWLSEVERE 2879
Matches	8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Best Local Similarity	88.4%; Score 38; DB 22; Length 5107;
8; Conservative	88.9%; Pred. NO. 3.5e+02;
1; Indels	
0; Gaps	
0;	
RESULT 4	
AAAG22041	
ID	AAAG22041 standard; Protein: 441 AA.
XX	
AC	AAAG22041;
XX	
DT	17-OCT-2000 (first entry)
XX	
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 24819.
KW	Protein identification; signal transduction pathway; metabolic pathway;
KW	hybridisation assay; genetic mapping; gene expression control; promoter;
XX	termination sequence.
XX	
OS	Arabidopsis thaliana.
XX	
PN	EP1033405-A2.
XX	
PD	06-SEP-2000.
XX	
PF	25-FEB-2000; 2000EP-0301439.
XX	
PR	25-FEB-1999; 99US-0121825.
PR	05-MAR-1999; 99US-0123180.
PR	09-MAR-1999; 99US-0123548.
PR	23-MAR-1999; 99US-0125788.
PR	25-MAR-1999; 99US-0126284.
PR	29-MAR-1999; 99US-0126785.
PR	01-APR-1999; 99US-0127462.
PR	06-APR-1999; 99US-0128234.
PR	08-APR-1999; 99US-0128714.
PR	16-APR-1999; 99US-0129845.
PR	19-APR-1999; 99US-0130077.
PR	21-APR-1999; 99US-0130449.
PR	23-APR-1999; 99US-0130510.
PR	28-APR-1999; 99US-0130891.
PR	30-APR-1999; 99US-0131449.
PR	30-APR-1999; 99US-0132048.
PR	04-MAY-1999; 99US-0132407.
PR	05-MAY-1999; 99US-0132484.
PR	06-MAY-1999; 99US-0132485.
PR	06-MAY-1999; 99US-0132486.
PR	07-MAY-1999; 99US-0132487.
PR	11-MAY-1999; 99US-0132863.
PR	14-MAY-1999; 99US-0134218.
PR	14-MAY-1999; 99US-0134219.
PR	14-MAY-1999; 99US-0134221.
PR	18-MAY-1999; 99US-0134370.
PR	19-MAY-1999; 99US-0134768.
PR	20-MAY-1999; 99US-0134941.
PR	99US-0135124.

PR	21-MAY-1999	9905-01353553
PR	24-MAY-1999	9905-01356621
PR	27-MAY-1999	9905-01360629
PR	28-MAY-1999	9905-01363682
PR	01-JUN-1999	9905-01372722
PR	03-JUN-1999	9905-01375728
PR	04-JUN-1999	9905-01375702
PR	07-JUN-1999	9905-01377724
PR	08-JUN-1999	9905-01380904
PR	10-JUN-1999	9905-01385400
PR	10-JUN-1999	9905-01388447
PR	14-JUN-1999	9905-01391119
PR	16-JUN-1999	9905-01394522
PR	16-JUN-1999	9905-01394653
PR	17-JUN-1999	9905-01394642
PR	18-JUN-1999	9905-01394654
PR	18-JUN-1999	9905-01394655
PR	18-JUN-1999	9905-01394558
PR	18-JUN-1999	9905-01394557
PR	18-JUN-1999	9905-01394577
PR	18-JUN-1999	9905-01394588
PR	18-JUN-1999	9905-01394595
PR	18-JUN-1999	9905-01394600
PR	18-JUN-1999	9905-01394611
PR	18-JUN-1999	9905-01394622
PR	18-JUN-1999	9905-01394633
PR	18-JUN-1999	9905-01394650
PR	18-JUN-1999	9905-01394753
PR	21-JUN-1999	9905-01396173
PR	21-JUN-1999	9905-01396177
PR	22-JUN-1999	9905-01396949
PR	23-JUN-1999	9905-01400353
PR	23-JUN-1999	9905-01400354
PR	24-JUN-1999	9905-01400695
PR	28-JUN-1999	9905-01400823
PR	29-JUN-1999	9905-01400921
PR	30-JUN-1999	9905-01412807
PR	01-JUL-1999	9905-01414842
PR	01-JUL-1999	9905-01421554
PR	02-JUL-1999	9905-01420055
PR	06-JUL-1999	9905-01423290
PR	08-JUL-1999	9905-01428203
PR	09-JUL-1999	9905-01428220
PR	12-JUL-1999	9905-01429277
PR	13-JUL-1999	9905-01433442
PR	14-JUL-1999	9905-01436524
PR	15-JUL-1999	9905-01440605
PR	16-JUL-1999	9905-01440855
PR	16-JUL-1999	9905-01440866
PR	19-JUL-1999	9905-01444205
PR	19-JUL-1999	9905-01444331
PR	19-JUL-1999	9905-01444332
PR	19-JUL-1999	9905-01444333
PR	19-JUL-1999	9905-01444334
PR	19-JUL-1999	9905-01444335
PR	20-JUL-1999	9905-01444352
PR	20-JUL-1999	9905-01446321
PR	22-JUL-1999	9905-01450879
PR	22-JUL-1999	9905-01450887
PR	22-JUL-1999	9905-01451512
PR	22-JUL-1999	9905-01451518
PR	23-JUL-1999	9905-01451518
PR	23-JUL-1999	9905-01451524
PR	26-JUL-1999	9905-01455276
PR	27-JUL-1999	9905-01455133
PR	27-JUL-1999	9905-01455138
PR	28-JUL-1999	9905-01455911
PR	02-AUG-1999	9905-01465862

PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 22-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0150566.
PR 25-AUG-1999; 99US-0150884.
PR 26-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159654.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160880.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.

PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 76.7%; Score 33; DB 21; Length 441;
Best Local Similarity 66.7%; Pred. No. 2.4e+02;
Matches 6; Conservativity 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 NMLSEVERE 9
Db 11 SMLDEIERE 19

RESULT 5
AAU72534
ID AAU72534 standard; Protein; 784 AA.

XX AAU72534;

DT 26-FEB-2002 (first entry)

DE Arabidopsis cell cycle protein CCP24.

XX Cell cycle protein; CCP; cell cycle regulation; herbicide;

KW plant growth regulator; plant development; abiotic stress; biotic stress;

KM nutrient deprivation; pathogen attack; crop yield; immunogen.

XX Arabidopsis thaliana.

OS Arabidopsis thaliana.

PN MO200185946-A2.

XX 15-NOV-2001.

PD 14-MAY-2001; 2001MO-IB01307.

PE 12-MAY-2000; 2000US-204045P.

PR (CROP-) CROPDSEIGN NV.

PA Inze D, Boudolf V, De Veylder L, Acosta JAT, Magyar Z;

XX WPI: 2002-062249/08.

DR N-PSDB; AAS96324.

PT New cell cycle protein and nucleic acid molecule encoding it useful for

XX regulating cell cycle progression in plants and for identifying

PS modulators which are useful as herbicides or plant growth regulators

XX Claim 34; Fig 26; 316pp; English.

CC The invention relates to a novel cell cycle protein (CCP) and the
CC polynucleotides encoding them. CCP is useful for identifying a compound
CC which modulates the activity of the polypeptide and which binds to the
CC polypeptide and an anti-CCP antibody is useful for detecting the presence
CC of CCP in a sample. A CCP modulator is useful for modulating the cell
CC cycle or growth of a plant such as Arabidopsis thaliana, rice, wheat,
CC maize, tomato, alfalfa, oilseed rape, soybean, sunflower and canola.
CC CCP nucleic acid and polypeptide molecules are useful as modulating
CC agents in regulating cell cycle progression in plants. CCP is useful to
CC treat disorders characterised by insufficient or excessive production of
CC CCP protein or production of CCP protein forms which have decreased or
CC aberrant activity. Compounds that bind to or modulate the activity
CC of CCP polypeptide are useful as herbicides or plant growth regulators.
CC The polynucleotide is useful for modifying cell fate, plant development,
CC plant morphology, biochemistry and/or physiology, the length of the G1,
CC S, G2 and/or M phase of the cell cycle of a plant, initiation, promotion,
CC stimulation or enhancement of cell division, DNA replication, seed set,
CC seed size, seed development, tuber, fruit, leaf formation, shoot and root
CC initiation and/or development, nodule function, dwarfism in plants,
CC senescence, tolerance or resistance to stress. CCP, the polynucleotide
CC and the anti-CCP antibody are useful in agriculture to modulate the
CC protein levels or activity of a protein involved in the cell cycle due

to environmental conditions, including abiotic stress such as cold, nutrient deprivation, heat, drought, salt stress, or biotic stress such as pathogen attack, to modulate e.g. enhance crop yields, and attenuate plant architecture, plant quality traits, plant reproduction and seed development, endoreduplication in storage cells, storage tissues and/or storage organs of plants or its parts. CCP is useful as an immunogen to generate antibodies. CCP protein is useful to screen for naturally occurring CCP substrates. The polynucleotide is useful for expressing CCP protein, to detect CCP mRNA, or a genetic lesion in a CCP gene and to modulate CCP activity. The present sequence represents a CCP protein of the invention.

Sequence 784 AA:

Query Match 76.7%; Score 33; DB 23; Length 784;
Best Local Similarity 66.7%; Pred. No. 4.4e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NMUSEVERE 9
: 11111111
Db 438 SMLEETERE 446

RESULT 6

AAV28897
ID AAV28897 standard; Protein; 172 AA.

AC AAV28897;

DT 25-JAN-2000 (first entry)

DE Human centrin-2 protein (Hcen-2).

Human centrin-2 protein; Hcen-2; neuropsychiatric disorder; marker; chromosome 18p; short arm; mutation; bipolar-affective disorder; BAD; identification; predisposition; Hcen-2 mediated disorder; primer; probe; DNA library; detection; mutant allele; therapeutic; modulate; screening; gene replacement therapy; antibody; compound; ameliorate; regulation; genetic counselling; diagnosis; gene expression; assay; treatment; genetic mapping; physical mapping.

OS Homo sapiens.

PN WO951186-A2.

PD 14-OCT-1999.

PF 05-APR-1999; 99WO-US07478.

PR 06-APR-1998; 98US-0080841.

PA (MILL-) MILLENNIUM PHARM INC.
(RECC) UNIV CALIFORNIA.

PI Chen H, Frelmer NB;

PI WPI; 1999-620160/53.

DR N-PSDB; AAX90997.

Novel methods and compositions for the diagnosis and treatment of neuropsychiatric disorders using compounds which modulate the Hcen-2 gene -
Claim 1; Fig 1; 78pp; English.

The present sequence is the human centrin-2 (Hcen-2) protein. The gene encoding this protein, is mapped to a portion on the short arm of chromosome 18 (18p). A genetic mutation in the wild type Hcen-2 sequence is involved in mediating neuropsychiatric disorders, such as bipolar-affective disorder (BAD). This sequence can be used to identify individuals with a predisposition to Hcen-2 mediated disorders. The Hcen-2 gene can be used as a marker for genetic and physical mapping of chromosome 18. It is a source of primers and probes, that can screen

DNA libraries for detection of mutant alleles that may be useful in therapeutic and diagnostic methods. This sequence may be utilized in gene replacement therapy, to generate antibodies, in diagnostic assays and for the identification of proteins, that may be involved in regulation of Hcen-2 mediated disorders. Compounds modulating expression of Hcen-2 gene are identified, that can be used to ameliorate Hcen-2 mediated disorders. This sequence is of great importance for genetic counselling, diagnosis and treatment of individuals predisposed to neuropsychiatric disorders.

Sequence 172 AA:

Query Match 74.4%; Score 32; DB 20; Length 172;
Best Local Similarity 75.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 MISEVDRE 9
: 11111111
Db 72 MISEVDRE 79

RESULT 7

ABG11803
ID ABG11803 standard; Protein; 215 AA.

AC ABG11803;

DT 13-FEB-2002 (first entry)

DE Novel human diagnostic protein #11794.

Human; chromosome mapping; gene mapping; gene therapy; forensic; food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

PA (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

PI WPI; 2001-639362/73.

DR N-PSDB; AAS75990.

New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity -
Claim 20; SEQ ID No 42162; 103pp; English.

The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, PCR primers, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations

CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG0010-ABG0377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 215 AA;

Query Match 74.4%; Score 32; DB 22; Length 215;
Best Local Similarity 66.7%; Pred. No. 1.7e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NMLSEVERE 9
Db 123 NLMSEVESE 131
:::|||||

RESULT 8
ABB71994
ID ABB71994 standard; Protein: 222 AA.
XX
AC ABB71994;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 42774.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
XX
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PMD, Myers EW;
XX
DR WPI: 2001-656860/75.
XX
DR N-PDB; ABL16097.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -
XX
PS Disclosure; SEQ ID NO 42774; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA
CC sequences (AB11840-AB116175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 222 AA;

QY 2 MLESEVERE 9
Db 112 MLEKIERE 119
||||:|||||

RESULT 9
AAU52422
ID AAU52422 standard; Protein: 415 AA.
XX
AC AAU52422;
XX
DT 27-FEB-2002 (first entry)
XX
DE Propionibacterium acnes immunogenic protein #13318.
XX
KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX
OS Propionibacterium acnes.
XX
PN WO200181581-A2.
XX
PD 01-NOV-2001.
XX
PF 20-APR-2001; 2001WO-US12865.
XX
PR 21-APR-2000; 2000US-199047P.
XX
PR 02-JUN-2000; 2000US-208841P.
XX
PR 07-JUL-2000; 2000US-216747P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;
XX
PI L'valsommeuve J, Zhang Y, Jen S, Carter D;
XX
DR WPI: 2001-616774/71.
XX
DR N-PDB; AAS59554.
XX
PT Propionibacterium acnes polypeptides and nucleic acids useful for
XX vaccinating against and diagnosing infections, especially useful for
XX treating acne vulgaris -
XX
PS Example 1; SEQ ID NO 13617; 1069pp; English.
XX
CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA).
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 415 AA;

Query Match 74.4%; Score 32; DB 22; Length 415;
Best Local Similarity 75.0%; Pred. No. 3.5e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NMLSEVER 8
1:1:111
Db 340 NMLSEVER 347

RESULT 10
AAG22042
ID AAG22042 standard; Protein; 430 AA.

AC AAG22042;

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 24820.

XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
KM termination sequence.

XX Arabidopsis thaliana.

OS EPI033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

PR 25-FEB-1999; 990S-0121825.

PR 05-MAR-1999; 990S-0123180.

PR 09-MAR-1999; 990S-0123548.

PR 23-MAR-1999; 990S-0125788.

PR 25-MAR-1999; 990S-0126264.

PR 29-MAR-1999; 990S-0126785.

PR 01-APR-1999; 990S-0127462.

PR 06-APR-1999; 990S-0128234.

PR 08-APR-1999; 990S-0128714.

PR 16-APR-1999; 990S-0129845.

PR 19-APR-1999; 990S-0130077.

PR 21-APR-1999; 990S-0130449.

PR 23-APR-1999; 990S-0130510.

PR 23-APR-1999; 990S-0130891.

PR 28-APR-1999; 990S-0131449.

PR 30-APR-1999; 990S-0132048.

PR 30-APR-1999; 990S-0132407.

PR 04-MAY-1999; 990S-0132484.

PR 05-MAY-1999; 990S-0132485.

PR 06-MAY-1999; 990S-0132486.

PR 06-MAY-1999; 990S-0132487.

PR 07-MAY-1999; 990S-0132863.

PR 11-MAY-1999; 990S-0134256.

PR 14-MAY-1999; 990S-0134216.

PR 14-MAY-1999; 990S-0134219.

PR 14-MAY-1999; 990S-0134221.

PR 14-MAY-1999; 990S-0134370.

PR 18-MAY-1999; 990S-0134768.

PR 19-MAY-1999; 990S-0134941.

PR 20-MAY-1999; 990S-0135124.

PR 21-MAY-1999; 990S-0135353.

PR 24-MAY-1999; 990S-0135629.

PR 25-MAY-1999; 990S-0136021.

PR 27-MAY-1999; 990S-0136392.

PR 28-MAY-1999; 990S-0136782.

PR 01-JUN-1999; 990S-0137222.

PR 03-JUN-1999; 990S-0137528.

PR 04-JUN-1999; 990S-0137502.

PR 07-JUN-1999; 990S-0137724.

PR 08-JUN-1999; 990S-0138094.

PR 10-JUN-1999; 990S-0138540.

PR 10-JUN-1999; 990S-0138847.

PR 14-JUN-1999; 990S-0139119.

PR 16-JUN-1999; 990S-0139452.

PR 16-JUN-1999; 990S-0139453.

PR 17-JUN-1999; 990S-0139492.

PR 18-JUN-1999; 990S-0139454.

PR 18-JUN-1999; 990S-0139455.

PR 18-JUN-1999; 990S-0139456.

PR 18-JUN-1999; 990S-0139457.

PR 18-JUN-1999; 990S-0139458.

PR 18-JUN-1999; 990S-0139459.

PR 18-JUN-1999; 990S-0139460.

PR 18-JUN-1999; 990S-0139461.

PR 18-JUN-1999; 990S-0139462.

PR 18-JUN-1999; 990S-0139463.

PR 18-JUN-1999; 990S-0139750.

PR 18-JUN-1999; 990S-0139763.

PR 21-JUN-1999; 990S-0139817.

PR 22-JUN-1999; 990S-0139899.

PR 23-JUN-1999; 990S-0140353.

PR 23-JUN-1999; 990S-0140354.

PR 24-JUN-1999; 990S-0140695.

PR 28-JUN-1999; 990S-0140823.

PR 29-JUN-1999; 990S-0140991.

PR 30-JUN-1999; 990S-0141842.

PR 01-JUL-1999; 990S-0142154.

PR 02-JUL-1999; 990S-0142055.

PR 06-JUL-1999; 990S-0142390.

PR 08-JUL-1999; 990S-0142803.

PR 09-JUL-1999; 990S-0142920.

PR 12-JUL-1999; 990S-0142977.

PR 13-JUL-1999; 990S-0143542.

PR 14-JUL-1999; 990S-0143624.

PR 15-JUL-1999; 990S-0144005.

PR 16-JUL-1999; 990S-0144085.

PR 16-JUL-1999; 990S-0144086.

PR 19-JUL-1999; 990S-0144325.

PR 19-JUL-1999; 990S-0144331.

PR 19-JUL-1999; 990S-0144332.

PR 19-JUL-1999; 990S-0144333.

PR 19-JUL-1999; 990S-0144334.

PR 20-JUL-1999; 990S-0144335.

PR 20-JUL-1999; 990S-0144632.

PR 20-JUL-1999; 990S-0144884.

PR 21-JUL-1999; 990S-0144814.

PR 21-JUL-1999; 990S-0145086.

PR 21-JUL-1999; 990S-0145088.

PR 22-JUL-1999; 990S-0145085.

PR 22-JUL-1999; 990S-0145087.

PR 22-JUL-1999; 990S-0145089.

PR 22-JUL-1999; 990S-0145192.

PR 23-JUL-1999; 990S-0145145.

PR 23-JUL-1999; 990S-0145218.

PR 23-JUL-1999; 990S-0145224.

PR 26-JUL-1999; 990S-0145276.

PR 27-JUL-1999; 990S-0145913.

PR 27-JUL-1999; 990S-0145918.

PR 27-JUL-1999; 990S-0145919.

PR 28-JUL-1999; 990S-0145951.

PR 02-AUG-1999; 990S-0146386.

PR 02-AUG-1999; 990S-0146388.

PR 02-AUG-1999; 990S-0146389.

PR 03-AUG-1999; 990S-0146389.

PR 04-AUG-1999; 990S-0147038.

PR 04-AUG-1999; 990S-0147204.

PR 04-AUG-1999; 990S-0147302.

PR 05-AUG-1999; 990S-0147192.

PR 05-AUG-1999; 990S-0147260.

PR 06-AUG-1999; 990S-0147303.

PR 06-AUG-1999; 990S-0147416.

PR 09-AUG-1999; 990S-0147493.

PR 09-AUG-1999; 990S-0147935.

PR 10-AUG-1999; 990S-0148171.

PR 11-AUG-1999; 990S-0148319.

PR 12-AUG-1999; 990S-0148341.

PR 13-AUG-1999; 990S-0148565.

```
PR 13-AUG-1999; 9905-0148684.
PR 16-AUG-1999; 9905-0149368.
PR 17-AUG-1999; 9905-0149175.
PR 18-AUG-1999; 9905-0149426.
PR 20-AUG-1999; 9905-0149722.
PR 20-AUG-1999; 9905-0149723.
PR 20-AUG-1999; 9905-0149929.
PR 23-AUG-1999; 9905-0149902.
PR 25-AUG-1999; 9905-0150566.
PR 26-AUG-1999; 9905-0150884.
PR 27-AUG-1999; 9905-0151065.
PR 27-AUG-1999; 9905-0151066.
PR 27-AUG-1999; 9905-0151080.
PR 30-AUG-1999; 9905-0151303.
PR 31-AUG-1999; 9905-0151438.
PR 01-SEP-1999; 9905-0151930.
PR 07-SEP-1999; 9905-0152363.
PR 10-SEP-1999; 9905-0153070.
PR 13-SEP-1999; 9905-0153758.
PR 15-SEP-1999; 9905-0154018.
PR 16-SEP-1999; 9905-0154039.
PR 22-SEP-1999; 9905-0154779.
PR 22-SEP-1999; 9905-0155139.
PR 23-SEP-1999; 9905-0153486.
PR 24-SEP-1999; 9905-0155659.
PR 28-SEP-1999; 9905-0156458.
PR 29-SEP-1999; 9905-0156596.
PR 04-OCT-1999; 9905-0157117.
PR 05-OCT-1999; 9905-0157753.
PR 06-OCT-1999; 9905-0157865.
PR 07-OCT-1999; 9905-0158029.
PR 08-OCT-1999; 9905-0158232.
PR 12-OCT-1999; 9905-0158369.
PR 13-OCT-1999; 9905-0159293.
PR 13-OCT-1999; 9905-0159294.
PR 13-OCT-1999; 9905-0159295.
PR 14-OCT-1999; 9905-0159329.
PR 14-OCT-1999; 9905-0159330.
PR 14-OCT-1999; 9905-0159331.
PR 14-OCT-1999; 9905-0159637.
PR 14-OCT-1999; 9905-0159638.
PR 18-OCT-1999; 9905-0159584.
PR 21-OCT-1999; 9905-0160741.
PR 21-OCT-1999; 9905-0160767.
PR 21-OCT-1999; 9905-0160768.
PR 21-OCT-1999; 9905-0160770.
PR 21-OCT-1999; 9905-0160814.
PR 21-OCT-1999; 9905-0160815.
PR 22-OCT-1999; 9905-0160980.
PR 22-OCT-1999; 9905-0160981.
PR 22-OCT-1999; 9905-0160989.
PR 25-OCT-1999; 9905-0161404.
PR 25-OCT-1999; 9905-0161405.
PR 25-OCT-1999; 9905-0161406.
PR 26-OCT-1999; 9905-0161359.
PR 26-OCT-1999; 9905-0161360.
PR 26-OCT-1999; 9905-0161361.
PR 28-OCT-1999; 9905-0161920.
PR 28-OCT-1999; 9905-0161922.
PR 28-OCT-1999; 9905-0161993.
PR 29-OCT-1999; 9905-0162142.
```

```
Query Match 74.4%; Score 32; DB 21; Length 430;
Best Local Similarity 75.0%; Pred. No. 3.6e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 2 MLEVERE 9
DB 1 MLEVERE 8
```

RESULT 11

```
ABBS3048
ID ABB53048 standard; Protein; 629 AA.
XX
AC ABB53048;
XX
DT 11-FEB-2002 (first entry)
XX
DE Escherichia coli polypeptide SEQ ID NO 1466.
XX
KW Escherichia coli; B2/D+A-; antiinflammatory; antibacterial;
KW immunosuppressive; extra-intestinal infection; phylogeny; meningitis;
KW systemic infection; non-diarrhoeal infection; septicaemia;
KW pyelonephritis; antibiotic resistance.
XX
OS Escherichia coli.
XX
PN WO200166572-A2.
XX
PD 13-SEP-2001.
XX
PF 12-MAR-2001; 2001WO-EP03445.
XX
PR 10-MAR-2000; 2000EP-0003145.
PR 02-FEB-2001; 2001EP-0001449.
XX
PA (INRM ) INSERM INST SANTE & RECH MEDICALE.
XX
PI Bingen E, Bonacorsi S, Clermont O, Nassif X, Tinsley C;
XX WPI; 2001-550253/61.
XX
PT A library of DNA fragments of Escherichia coli strains for the
PT phylogenetic determination of a given strain comprises polynucleotides of
PT nature B2/D+ A- -
XX
PS Example 6; Fig 6; 646pp; English.
XX
CC The invention relates to a library of DNA fragments of Escherichia coli
CC strains comprising polynucleotides (ABA88577-ABA88729 and ABA89533)
CC and encoded proteins (ABBS2459-ABBS2919 and ABBS2954-ABBS3094) of nature
CC B2/D+A-. The polynucleotides have potential antiinflammatory,
CC antibacterial and immunosuppressive activity as part of pharmaceutical
CC compositions used to treat, palliate or prevent extra-intestinal E. coli
CC infections. The polypeptides are useful for determining the phylogenetic
CC group of a given E. coli strain. These polypeptides can detect and treat
CC an undesired development of E. coli, particularly an extra-intestinal
CC infection that include systemic and non-diarrhoeal infections such as
CC septicaemia, pyelonephritis and meningitis this is particularly
CC advantageous as bacterial resistance is increasing with the more
CC frequent use of broad spectrum antibiotics.
XX
SQ Sequence 629 AA;
```

```
Query Match 74.4%; Score 32; DB 22; Length 629;
Best Local Similarity 66.7%; Pred. No. 5.5e+02;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 NMLEVERE 9
DB 478 SMLTELERE 486
```

RESULT 12

```
ABG09932
ID ABG09932 standard; Protein; 692 AA.
```

```
AC ABG09932;
```

```
DT 13-FEB-2002 (first entry)
```

```
DE Novel human diagnostic protein #9923.
```

```
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
```

KM food supplement; medical imaging; diagnostic; genetic disorder.
 XX Homo sapiens.
 OS
 XX WO200175067-A2.
 PN
 XX 11-OCT-2001.
 PD
 XX 30-MAR-2001; 2001WO-US08631.
 PF
 XX 31-MAR-2000; 2000US-0540217.
 PR
 XX 23-AUG-2000; 2000US-0649167.
 PA
 XX (HSE-) HSEQ INC.
 PI
 XX Drmanac RT, Liu C, Tang YT;
 DR
 XX MPI: 2001-639362/73.
 DR
 XX N-PSDB: AAS74119.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 PS
 XX Claim 20; SEQ ID NO 40291; 103pp; English.
 PS
 XX The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations in
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX
 SQ Sequence 692 AA:
 Query Match 74.4%; Score 32; DB 22; Length 692;
 Best Local Similarity 77.8%; Pred. No. 6.1e+02;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 OY 1 NMLSEVERE 9
 Db 493 NDLTEVERE 501
 RESULT 13
 AAB27497
 ID AAB27497 standard; Protein; 930 AA.
 AC AAB27497;
 XX
 XX 02-JUL-2002 (first entry)
 DT
 XX Streptococcus polypeptide SEQ ID NO 4170.
 DE
 XX Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
 KW group A streptococcus; Streptococcus pyogenes; antibacterial;
 KW antiinflammatory; infection; vaccine; meningitis; gene therapy.
 KM

XX Streptococcus agalactiae.
 OS
 XX WO200234771-A2.
 PN
 XX 02-MAY-2002.
 PD
 XX 29-OCT-2001; 2001WO-GB04789.
 PF
 XX 27-OCT-2000; 2000GB-0026333.
 PR
 XX 24-NOV-2000; 2000GB-0028727.
 PR
 XX 07-MAR-2001; 2001GB-0005640.
 PA
 XX (CHIR-) CHIRON SPA.
 PA
 XX (GENO-) INST GENOMIC RES.
 PI
 XX Telford J, Maignani V, Margarit Ros YI, Grandi G, Fraser C;
 PI
 XX Tettelein H;
 DR
 XX MPI: 2002-352536/38.
 DR
 XX N-PSDB: ABN68128.
 XX
 PT New Streptococcus protein for the treatment or prevention of infection
 PT or disease caused by Streptococcus bacteria, such as meningitis, and
 PT for detecting a compound that binds to the protein -
 PS
 XX Claim 1; Page 3573; 4525pp; English.
 PS
 XX The invention relates to a protein (AAB25413-ABP30895) from group B
 CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
 CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
 CC the specification. The proteins have antibacterial and antiinflammatory
 CC activity. (II), nucleic acids encoding (I), ABN66044-ABN71526 and
 CC antibodies that bind (I) are used in the manufacture of medicaments for
 CC the treatment or prevention of infection or disease caused by
 CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
 CC Nucleic acids encoding (I) are used to detect Streptococcus in a
 CC biological sample. (I) is used to determine whether a compound binds to
 CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
 CC used as a vaccine or diagnostic composition. The disease caused by
 CC Streptococcus that is prevented or treated may be meningitis. Nucleic
 CC acid encoding (I) may be used to recombinantly produce (I) and may be
 CC used in gene therapy. Antibodies to (I) are used for affinity
 CC chromatography, immunoassays, and distinguishing/identifying
 CC Streptococcus proteins.
 CC
 XX
 SQ Sequence 930 AA:
 Query Match 74.4%; Score 32; DB 23; Length 930;
 Best Local Similarity 66.7%; Pred. No. 8.3e+02;
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 OY 1 NMLSEVERE 9
 Db 48 NDLSEVKE 56
 RESULT 14
 AAB60483
 ID AAB60483 standard; Protein; 933 AA.
 AC AAB60483;
 XX
 XX 24-APR-2001 (first entry)
 DT
 XX Human cell cycle and proliferation protein CCYR-31, SEQ ID NO:31.
 DE
 XX Cell cycle and proliferation protein; CCYR; human; agonist;
 KW antagonist; gene therapy; detection; gene therapy;
 KW transgenic animal disease model; immune disorder;
 KW developmental disorder; cell signalling disorder;
 KW cell proliferative disorder; cancer; tumour; anaemia; epilepsy;
 KW arteriosclerosis; asthma; allergy; diabetes mellitus;
 KM

KW	menstrual cycle disorder; bacterial infection.
XX	
OS	Homo sapiens.
XX	
PN	WO200107471-A2.
XX	
PD	01-FEB-2001.
XX	
PF	21-JUL-2000; 2000WO-US19948.
XX	
PR	21-JUL-1999; 99US-0145075.
PR	08-SEP-1999; 99US-0153129.
PR	10-NOV-1999; 99US-0164647.
XX	
PA	(INCY-) INCYTE GENOMICS INC.
XX	
PI	Hillman JL, Lal P, Tang YT, Yue H, Au-Young J, Bandman O;
PI	Azlmzal Y, Yang J, Lu DAM, Baughn MR, Patterson C, Shah P;
XX	
DR	WPI: 2001-112727/12.
XX	
DR	N-PSDB; AAF59620.
XX	
PT	Human cell cycle and proliferation proteins and polynucleotides are
PT	used to treat, diagnose and prevent immune, developmental and cell
PT	signaling disorders and cell proliferative disorders including cancer -
XX	
PS	Claim 1; Page 136-138; 205pp; English.

CC Sequences AAB604533-AAB60506 represent 54 human cell cycle and
CC proliferation proteins (CCYPR), which are encoded by AAF59590-AAF59643.
CC CCYPR and agonists of CCYPR are used to treat diseases or conditions
CC associated with decreased expression of functional CCYPR, while CCYPR
CC antagonists are used to treat diseases or conditions associated with
CC overexpression of functional CCYPR. Monoclonal or polyclonal antibodies
CC to CCYPR may be used in enzyme-linked immunosorbent assays (ELISA) or
CC radioimmunoassays to detect CCYPR. CCYPR itself may be used to detect
CC compounds e.g., antibodies, oligonucleotides and proteins (receptors)
CC that specifically bind to CCYPR, and in drug screening methods to
CC identify compounds that modulate the activity of CCYPR. CCYPR
CC nucleotides can be used to generate transgenic animal models of human
CC disease, and can be used in gene therapy in target cells with genetic
CC abnormalities with respect to the expression of CCYPR for the
CC treatment or prevention of a disorder associated with CCYPR.
CC Diseases which can be diagnosed, treated and prevented using CCYPR
CC proteins, nucleic acids, agonists or antagonists include immune,
CC developmental and cell signalling disorders, and cell proliferative
CC disorders including cancer. Specific examples of these disorders
CC include anaemia, epilepsy, arteriosclerosis, asthma, cancer, allergies,
CC diabetes mellitus, disorders of the menstrual cycle and infections
CC caused by bacteria.

Seq	sequence	933 AA:
Query Match	74.4%	Score 32; DB 22;
Best Local Similarity	55.6%;	Length 933;
Matches	5; Conservative	Pred. No. 8.4e+02;
		Indels 0; Gaps 0;

```

      |::|||::
Db    343 NLISEVEKD 351

```

RESULT	15
ABG09933	
ID	ABG09933 standard; Protein; 1798 AA.
XX	
XX	
AC	ABG09933:
XX	
DT	13-FEB-2002 (first entry)
XX	
DE	Novel human diagnostic protein #9924.
XX	
FW	Human; chromosome mapping; gene mapping; gene therapy; forensic;

XX	food supplement; medical imaging; diagnostic; genetic disorder.
XX	
OS	Homo sapiens.
PN	WO200175067-A2.
XX	
PD	11-OCT-2001.
XX	
PF	30-MAR-2001; 2001MO-US08631.
XX	
PR	31-MAR-2000; 2000US-0540217.
XX	
PR	23-AUG-2000; 2000US-0649167.
XX	
PA	(HYSE-) HYSEQ INC.
XX	
PI	Drmanac RT, Liu C, Tang YT;
XX	
DR	WPI: 2001-639362/73.
XX	
DR	N-PSDB: AAS74120.
XX	
PT	New isolated polynucleotide and encoded polypeptides, useful in
PT	diagnostics, forensics, gene mapping, identification of mutations
PT	responsible for genetic disorders or other traits and to assess
PT	biodiversity -
XX	
PS	Claim 20; SEQ ID No 40292; 103bp; English.

CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (III). (II) is useful for generating antibodies against it, detecting or
CC quantifying a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

Sequence	1798 AA;
Query Match	74.4%;
Best Local Similarity	77.8%;
Matches	7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Score	32; DB 22; Length 1798;
Pred. No.	1.7e+03;

Db 482 NDJTEVERE 490

Search completed: November 18, 2002, 15:55:24
Job time : 28.129 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 18, 2002, 15:53:18 ; Search time 10.4516 Seconds
(without alignments)
25.336 Million cell updates/sec

Title: US-09-848-781-1

Perfect score: 43

Sequence: 1 NMLSEVERE 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA:*

1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCRTUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfilest1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	32	74.4	172	4	US-09-285-601-2
2	30	67.8	11	4	US-09-537-817B-8
3	30	67.8	19	4	US-09-537-817B-10
4	30	67.8	123	2	US-08-822-573-2
5	30	67.8	357	2	US-09-123-851-4
6	30	67.8	357	2	US-08-728-520-4
7	30	67.8	840	2	US-08-467-822-25
8	30	67.8	840	4	US-08-432-697-25
9	30	67.8	840	4	US-08-466-248-25
10	30	67.8	3169	4	US-09-453-702B-257
11	29	67.4	61	3	US-08-743-975-9
12	29	67.4	119	2	US-08-710-749-27
13	29	67.4	144	4	US-09-250-609-2
14	29	67.4	204	2	US-08-715-204-1
15	29	67.4	204	4	US-08-691-814B-10
16	29	67.4	204	3	US-09-162-597-1
17	29	67.4	204	4	US-09-250-609-9
18	29	67.4	204	4	US-09-250-609-9
19	29	67.4	219	4	US-08-928-213B-60
20	29	67.4	285	4	US-09-173-300-24
21	29	67.4	375	4	US-08-193-705-3
22	29	67.4	377	4	US-08-193-705-3
23	29	67.4	410	1	US-07-955-905A-26
24	29	67.4	591	2	US-08-937-540-4
25	29	67.4	591	4	US-09-398-395A-24
26	29	67.4	1077	4	US-09-390-234-12
27	29	67.4	2237	1	US-08-354-973-1

28	29	67.4	2516	3	US-08-374-077C-2	Sequence 2, Appli
29	29	67.4	2516	4	US-08-895-590-2	Sequence 2, Appli
30	29	67.4	2516	4	US-09-539-879A-2	Sequence 2, Appli
31	28	65.1	19	4	US-09-150-797-5	Sequence 5, Appli
32	28	65.1	92	4	US-09-134-001C-4014	Sequence 4014, Ap
33	28	65.1	95	4	US-08-858-207A-516	Sequence 516, App
34	28	65.1	117	4	US-09-134-001C-4706	Sequence 4706, Ap
35	28	65.1	129	4	US-08-942-088-4	Sequence 4, Appli
36	28	65.1	181	4	US-08-961-083-62	Sequence 62, Appli
37	28	65.1	245	4	US-08-482-918-63	Sequence 63, Appli
38	28	65.1	245	4	US-09-224-681-63	Sequence 63, Appli
39	28	65.1	245	4	US-08-336-728A-63	Sequence 63, Appli
40	28	65.1	248	2	US-08-955-848A-82	Sequence 82, Appli
41	28	65.1	266	4	US-08-482-918-57	Sequence 57, Appli
42	28	65.1	266	4	US-09-224-681-57	Sequence 57, Appli
43	28	65.1	266	4	US-08-336-728A-57	Sequence 57, Appli
44	28	65.1	271	4	US-08-482-918-52	Sequence 52, Appli
45	28	65.1	271	4	US-09-224-681-52	Sequence 52, Appli

ALIGNMENTS

```
RESULT 1
US-09-285-601-2
; Sequence 2, Application US/09285601
; Patent No. 6248528
; GENERAL INFORMATION:
; APPLICANT: Chen, Hong
; APPLICANT: Freimer, Nelson
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND
; TITLE OF INVENTION: TREATMENT OF NEUROPSYCHIATRIC DISORDERS
; FILE REFERENCE: 7853-089
; CURRENT APPLICATION NUMBER: US/09/285,601
; CURRENT FILING DATE: 1999-04-02
; EARLIER APPLICATION NUMBER: 60/080,841
; EARLIER FILING DATE: 1998-04-06
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 172
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-285-601-2

Query Match      74.4% Score 32; DB 4; Length 172;
Best Local Similarity 75.0%; Pred. No. 28;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      2 MLSEVERE 9
       1:|||||
DB      72 MISEVDRE 79

RESULT 2
US-09-537-817B-8
; Sequence 8, Application US/09537817B
; Patent No. 6417327
; GENERAL INFORMATION:
; APPLICANT: Shuichi Oka
; APPLICANT: Yoshimitsu Yamazaki
; APPLICANT: Toru Imamura
; APPLICANT: Yasuko Fujita
; APPLICANT: Saori Yamamoto
; APPLICANT: Yukiko Okita
; APPLICANT: Kazuo Ozawa
; APPLICANT: Reiko Akakura
; APPLICANT: Chikako Ito
; TITLE OF INVENTION: PEPTIDE CAPABLE OF REGULATING
; TITLE OF INVENTION: PHYSIOLOGICAL FUNCTION OF FGF-5 AND PHARMACEUTICAL
; TITLE OF INVENTION: COMPOSITION CONTAINING THE PEPTIDE
; FILE REFERENCE: 08206-012001
; CURRENT APPLICATION NUMBER: US/09/537,817B
```

; CURRENT FILING DATE: 2001-08-23
; PRIOR APPLICATION NUMBER: JP 88364/1999
; PRIOR FILING DATE: 1999-03-30
; PRIOR APPLICATION NUMBER: JP 6643/2000
; PRIOR FILING DATE: 2000-01-14
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: SEQ ID NO:8 represents a partial sequence
; OTHER INFORMATION: (corresponding to amino acids 113-123) of the
; OTHER INFORMATION: amino acid sequence of human FGF-5.
US-09-537-817B-8

Query Match 69.8%; Score 30; DB 4; Length 11;
Best Local Similarity 75.0%; Pred. No. 3.5;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 NMLSEVER 8
|1111:11|
Db 4 NMLSOVHR 11

RESULT 3
US-09-537-817B-10
; Sequence 10, Application US/09537817B
; Patent No. 6417327
; GENERAL INFORMATION:
; APPLICANT: Shuichi Oka
; APPLICANT: Yoshimitsu Yamazaki
; APPLICANT: Toru Imanura
; APPLICANT: Yasuko Fujita
; APPLICANT: Saori Yamamoto
; APPLICANT: Yukiko Orita
; APPLICANT: Kazuo Ozawa
; APPLICANT: Reiko Akakura
; APPLICANT: Chikako Ito
; TITLE OF INVENTION: PEPTIDE CAPABLE OF REGULATING
; TITLE OF INVENTION: PHYSIOLOGICAL FUNCTION OF FGF-5 AND PHARMACEUTICAL
; TITLE OF INVENTION: COMPOSITION CONTAINING THE PEPTIDE
; FILE REFERENCE: 08206-012001
; CURRENT APPLICATION NUMBER: US/09/537,817B
; CURRENT FILING DATE: 2001-08-23
; PRIOR APPLICATION NUMBER: JP 88364/1999
; PRIOR FILING DATE: 1999-03-30
; PRIOR APPLICATION NUMBER: JP 6643/2000
; PRIOR FILING DATE: 2000-01-14
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: SEQ ID NO:10 represents a partial sequence of the
; OTHER INFORMATION: amino acid sequence of human FGF-5.
US-09-537-817B-10

Query Match 69.8%; Score 30; DB 4; Length 19;
Best Local Similarity 75.0%; Pred. No. 6.4;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 NMLSEVER 8
|1111:11|
Db 12 NMLSOVHR 19

RESULT 4
US-08-822-573-2
; Sequence 2, Application US/08822573

; Patent No. 5874254
; GENERAL INFORMATION:
; APPLICANT: Imamura et al.
; TITLE OF INVENTION: FGF-5 ANALOGOUS PROTEIN, AND
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION
; TITLE OF INVENTION: CONTAINING SAME
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112-0228
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/822,573
; FILING DATE: March 19, 1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: Japanese Appln. No. 5874254 75994/1996
; FILING DATE: March 29, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Rochelle K. Seide
; REGISTRATION NUMBER: 32,300
; REFERENCE/DOCKET NUMBER: A30981
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 408-2500
; TELEFAX: (212) 765-2519
; TELEEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 123 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: Internal
US-08-822-573-2

Query Match 69.8%; Score 30; DB 2; Length 123;
Best Local Similarity 75.0%; Pred. No. 49;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 NMLSEVER 8
|1111:11|
Db 116 NMLSOVHR 123

RESULT 5
US-09-123-851-4
; Sequence 4, Application US/09123851
; Patent No. 5958405
; GENERAL INFORMATION:
; APPLICANT: Goli, Surya K.
; TITLE OF INVENTION: A NOVEL HUMAN PROTEIN TYROSINE KINASE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: U.S.
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5


```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/123,851
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/728,520
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0136 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 357 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 1166579
US-09-123-851-4

Query Match
Best Local Similarity 69.8%; Score 30; DB 2; Length 357;
Pred. No. 1.6e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 3 LSEVERE 9
DB 143 MSEVERE 149

RESULT 6
US-08-728-520-4
Sequence 4, Application US/08728520
Patent No. 5994112
GENERAL INFORMATION:
APPLICANT: Goll, Surya K.
TITLE OF INVENTION: A NOVEL HUMAN PROTEIN TYROSINE KINASE
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: U.S.
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/728,520
FILING DATE: Filed Herewith
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0136 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 357 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

```

```

MOLECULE TYPE: peptide
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 1166579
US-08-728-520-4

Query Match
Best Local Similarity 69.8%; Score 30; DB 2; Length 357;
Pred. No. 1.6e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 3 LSEVERE 9
DB 143 MSEVERE 149

RESULT 7
US-08-467-822-25
Sequence 25, Application US/08467822
Patent No. 5843460
GENERAL INFORMATION:
APPLICANT: Labigne, Agnes
APPLICANT: Sauterbaum, Sebastien
APPLICANT: Ferrero, Richard L.
APPLICANT: Thibierge, Jean-Michel
TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST
TITLE OF INVENTION: HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE
TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID
NUMBER OF SEQUENCES: 44
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flanagan, Henderson, Farabow, Garrett &
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/467,822
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/447,177
FILING DATE: 19-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/432,697
FILING DATE: 02-MAY-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Meyers, Kenneth J.
REGISTRATION NUMBER: 25,146
REFERENCE/DOCKET NUMBER: 03495.0137-02000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 408-4000
TELEFAX: (202) 408-4400
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 840 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-467-822-25

Query Match
Best Local Similarity 69.8%; Score 30; DB 2; Length 840;
Pred. No. 4e+02;
Matches 4; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 NMLSEVERE 9
1:1:1:1:
Db 297 NLLAEIKD 305

RESULT 8

US-08-432-697-25
; Sequence 25, Application US/08432697
; Patent No. 6248330
; GENERAL INFORMATION:
; APPLICANT: Labigne, Agnes
; APPLICANT: Sauerbaum, Sebastien
; APPLICANT: Ferrero, Richard L.
; APPLICANT: Thiberge, Jean-Michel
; TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST
; TITLE OF INVENTION: HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE
; TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID
; TITLE OF INVENTION: POLYPEPTIDES
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flinnegan, Henderson, Farabow, Garrett &
; ADDRESS: Dunnet
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/432,697
; FILING DATE: 02-MAY-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 03495.0137-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4400
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 840 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-432-697-25

Query Match 69.8%; Score 30; DB 4; Length 840;
Best Local Similarity 44.4%; Pred. No. 4e+02;

Matches 4; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 NMLSEVERE 9
1:1:1:1:
Db 297 NLLAEIKD 305

RESULT 9

US-08-466-248-25
; Sequence 25, Application US/08466248
; Patent No. 6258359
; GENERAL INFORMATION:
; APPLICANT: Labigne, Agnes
; APPLICANT: Sauerbaum, Sebastien
; APPLICANT: Ferrero, Richard L.
; APPLICANT: Thiberge, Jean-Michel
; TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST
; TITLE OF INVENTION: HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE

; TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID
; TITLE OF INVENTION: POLYPEPTIDES
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flinnegan, Henderson, Farabow, Garrett &
; ADDRESS: Dunnet
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,248
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/447,177
; FILING DATE: 19-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/432,697
; FILING DATE: 02-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 03495.0137-02000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4400
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 840 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-466-248-25

Query Match 69.8%; Score 30; DB 4; Length 840;
Best Local Similarity 44.4%; Pred. No. 4e+02;

Matches 4; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 NMLSEVERE 9
1:1:1:1:
Db 297 NLLAEIKD 305

RESULT 10

US-09-453-702B-257
; Sequence 257, Application US/09453702B
; Patent No. 6365723
; GENERAL INFORMATION:
; APPLICANT: Blattner, Frederick R.
; APPLICANT: Burland, Nicole T.
; APPLICANT: Perna, Nicole T.
; APPLICANT: Plunkett, Guy
; APPLICANT: Welch, Rod
; TITLE OF INVENTION: No. 6365723el Sequences of E. coli 0157
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53701-2113
; COMPUTER READABLE FORM:

```

; MEDIUM TYPE: Diskette, 3.50 inch, 1.44mb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 8.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/453,702B
; FILING DATE: 03-Dec-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/110,955
; FILING DATE: 04-DEC-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J.
; REGISTRATION NUMBER: 27386
; REFERENCE/DOCKET NUMBER: 960296.95017
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (608) 251-5000
; TELEFAX: (608) 251-9166
; INFORMATION FOR SEQ ID NO: 257:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3169 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 257:
US-09-453-702B-257

```

```

Query Match          69.8%; Score 30; DB 4; Length 3169;
Best Local Similarity 62.5%; Pred. No. 1.7e+03;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 NMLSEVER 8
Db 573 NMLSEIK 580

```

RESULT 11
US-08-743-975-9
; Sequence 9, Application US/08743975
; Patent No. 6057434
; GENERAL INFORMATION:
; APPLICANT: Jian Ni, Guo-Liang Yu and Reiner Gentz
; TITLE OF INVENTION: Mammary Transforming Protein
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CECCHI, BYRNE, BAIN, GILFILLAN,
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/743,975
; FILING DATE: 01 NOVEMBER 1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/006,187
; FILING DATE: 02 NOVEMBER 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: MOLLINS, J.G.
; REGISTRATION NUMBER: 33,073
; REFERENCE/DOCKET NUMBER: 325800-507 (PF212)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 9:

```

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 61 AMINO ACIDS
; TYPE: AMINO ACID
; STRANDEDNESS:
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PROTEIN
; US-08-743-975-9

```

```

Query Match          67.4%; Score 29; DB 3; Length 61;
Best Local Similarity 66.7%; Pred. No. 36;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

OY 1 NMLSEVER 9
Db 19 NMLSEPER 27

```

RESULT 12
US-08-710-749-27
; Sequence 27, Application US/08710749
; Patent No. 5955089
; GENERAL INFORMATION:
; APPLICANT: Brilles, David E.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Becker, Robert
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/710,749
; FILING DATE: 20-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2074
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 119 amino acids
; TYPE: amino acid
; STRANDEDNESS: n/a
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
; US-08-710-749-27

```

```

Query Match          67.4%; Score 29; DB 2; Length 119;
Best Local Similarity 55.6%; Pred. No. 75;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

```

OY 1 NMLSEVER 9
Db 47 NOVSEIDRE 55

```

RESULT 13
US-09-250-609-2
; Sequence 2, Application US/09250609A
; Patent No. 6458943

```

```

: GENERAL INFORMATION:
: APPLICANT: Byrne, Jennifer A.
: TITLE OF INVENTION: Members of the D52 Gene Family
: FILE REFERENCE: 1383.0210002
: CURRENT APPLICATION NUMBER: US/09/250.609A
: CURRENT FILING DATE: 1999-02-17
: NUMBER OF SEQ ID NOS: 108
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 2
: LENGTH: 144
: TYPE: PRT
: ORGANISM: Homo sapiens
: US-09-250-609-2

Query Match          67.4%; Score 29; DB 4; Length 144;
Best Local Similarity 66.7%; Pred. No. 92;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      1 NMLSEVERE 9
       :|||||:
Db      42 SMLSEEKE 50

RESULT 14
US-08-715-204-1
: Sequence 1, Application US/08715204
: Patent No. 5874286
: GENERAL INFORMATION:
: APPLICANT: Bandman, Olga
: APPLICANT: Au-Young, Janice
: APPLICANT: Goll, Surya K.
: APPLICANT: Hillman, Jennifer.
: APPLICANT: Zweigler, Gary B.
: TITLE OF INVENTION: A NOVEL TUMOR PROTEIN
: NUMBER OF SEQUENCES: 7
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Incyte Pharmaceuticals, Inc.
: STREET: 3174 Porter Drive
: CITY: Palo Alto
: STATE: CA
: COUNTRY: U.S.
: ZIP: 94304
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FastSeq Version 1.5
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/715,204
: FILING DATE: Filed Herewith
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Billings, Lucy J.
: REGISTRATION NUMBER: 36,749
: REFERENCE/DOCKET NUMBER: PF-0126 US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415-855-0555
: TELEFAX: 415-845-4166
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 204 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
: IMMEDIATE SOURCE:
: LIBRARY:
: CLONE: Consensus
: US-08-715-204-1

Query Match          67.4%; Score 29; DB 2; Length 204;
```

```

Best Local Similarity 66.7%; Pred. No. 1.4e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      1 NMLSEVERE 9
       :|||||:
Db      29 SMLSEEKE 37

RESULT 15
US-08-691-814B-10
: Sequence 10, Application US/08691814B
: Patent No. 5981218
: GENERAL INFORMATION:
: APPLICANT: Rio, Marie-Christine
: APPLICANT: Tomasello, Catherine
: APPLICANT: Basset, Paul
: APPLICANT: Byrne, Jennifer
: TITLE OF INVENTION: Isolated Nucleic Acid Molecules Useful
: as Leukemia Markers and in Breast Cancer Prognosis
: NUMBER OF SEQUENCES: 124
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
: STREET: 1100 New York Ave, NW, Suite 600
: CITY: Washington
: STATE: DC
: COUNTRY: USA
: ZIP: 20005-3934
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/691,814B
: FILING DATE: 31-JUL-1996
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 60/002,183
: FILING DATE: 09-AUG-1995
: ATTORNEY/AGENT INFORMATION:
: NAME: Steffe, Eric K.
: REGISTRATION NUMBER: 36,688
: REFERENCE/DOCKET NUMBER: 1383.0090001
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 202-371-2600
: TELEFAX: 202-371-2543
: INFORMATION FOR SEQ ID NO: 10:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 204 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-691-814B-10

Query Match          67.4%; Score 29; DB 2; Length 204;
Best Local Similarity 66.7%; Pred. No. 1.4e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      1 NMLSEVERE 9
       :|||||:
Db      29 SMLSEEKE 37
```

Search completed: November 18, 2002, 15:56:49
Job time : 12.4516 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 18, 2002, 15:56:17 ; Search time 4.93548 Seconds
(without alignments)
27.464 Million cell updates/sec

Title: US-09-848-781-1
Perfect score: 43
Sequence: 1 NMLSEVERE 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 97044 seqs, 15060890 residues

Total number of hits satisfying chosen parameters: 97044

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Database : Published Applications, AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the distribution, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	31	72.1	629	10	US-09-815-242-10961
2	30	69.8	30	10	US-09-864-761-38335
3	30	69.8	34	10	US-09-864-761-47991
4	30	69.8	185	10	US-09-815-242-11423
5	30	69.8	584	10	US-09-815-242-5013
6	30	69.8	589	10	US-09-815-242-10878
7	30	69.8	1294	10	US-09-815-242-13724
8	29	67.4	61	10	US-09-263-811-9
9	29	67.4	144	10	US-09-250-611-2
10	29	67.4	204	10	US-09-250-611-4
11	29	67.4	204	10	US-09-250-611-9
12	29	67.4	285	12	US-10-027-450-24
13	29	67.4	376	9	US-09-978-756-2
14	29	67.4	419	9	US-09-895-913A-30
15	29	67.4	491	9	US-09-895-913A-242
16	29	67.4	519	10	US-09-887-586A-24
17	29	67.4	591	10	US-09-903-012-24
18	28	65.1	78	10	US-09-815-242-4960
19	28	65.1	101	10	US-09-815-242-10804

20	28	65.1	139	10	US-09-925-300-1311	Sequence 1311, Ap
21	28	65.1	172	12	US-10-109-885-4	Sequence 4, Appl
22	28	65.1	181	10	US-09-765-272-62	Sequence 62, Appl
23	28	65.1	245	10	US-09-005-243-63	Sequence 63, Appl
24	28	65.1	245	10	US-09-224-683-63	Sequence 63, Appl
25	28	65.1	266	10	US-09-005-243-57	Sequence 57, Appl
26	28	65.1	266	10	US-09-224-683-57	Sequence 57, Appl
27	28	65.1	271	10	US-09-005-243-52	Sequence 52, Appl
28	28	65.1	271	10	US-09-224-683-52	Sequence 52, Appl
29	28	65.1	273	10	US-09-005-243-48	Sequence 48, Appl
30	28	65.1	273	10	US-09-005-243-50	Sequence 48, Appl
31	28	65.1	273	10	US-09-005-243-50	Sequence 49, Appl
32	28	65.1	273	10	US-09-005-243-53	Sequence 53, Appl
33	28	65.1	273	10	US-09-005-243-61	Sequence 53, Appl
34	28	65.1	273	10	US-09-224-683-48	Sequence 48, Appl
35	28	65.1	273	10	US-09-224-683-49	Sequence 49, Appl
36	28	65.1	273	10	US-09-224-683-50	Sequence 50, Appl
37	28	65.1	273	10	US-09-224-683-53	Sequence 53, Appl
38	28	65.1	273	10	US-09-224-683-61	Sequence 61, Appl
39	28	65.1	274	10	US-09-005-243-51	Sequence 51, Appl
40	28	65.1	274	10	US-09-224-683-51	Sequence 51, Appl
41	28	65.1	319	9	US-10-032-159A-12	Sequence 12, Appl
42	28	65.1	398	10	US-09-741-669-430	Sequence 430, App
43	28	65.1	559	10	US-09-738-973-396	Sequence 396, App
44	28	65.1	670	12	US-10-063-763-10	Sequence 10, Appl
45	28	65.1	1247	9	US-10-032-159A-8	Sequence 8, Appl

ALIGNMENTS

RESULT 1
US-09-815-242-10961
; Sequence 10961, Application US/09815242
; Patent No. US2002061569A1
; GENERAL INFORMATION:
; APPLICANT: Haseldeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA 011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10961
; LENGTH: 629
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-09-815-242-10961
Query Match 72.1%; Score 31; DB 10; Length 629;
Best Local Similarity 75.0%; Pred. No. 57;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 NMLSEVER 8
:|||||
Db 596 NMLSEVER 603

RESULT 2
US-09-864-761-38335
; Sequence 38335, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 38335
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC010974.3
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1
US-09-864-761-38335

Query Match 69.8%; Score 30; DB 10; Length 30;
Best Local Similarity 75.0%; Pred. No. 3;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 NMLSEVER 8
:|||||
Db 9 DMLSEVER 16

RESULT 3
US-09-864-761-47991
; Sequence 47991, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 47991
; LENGTH: 34
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC010974.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.89
US-09-864-761-47991

Query Match 69.8%; Score 30; DB 10; Length 34;
Best Local Similarity 75.0%; Pred. No. 3.5;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 NMLSEVER 8
:|||||
Db 13 DMLSEVER 20

```
RESULT 4
US-09-815-242-11423
; Sequence 11423, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11423
; LENGTH: 185
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-09-815-242-11423

Query Match          69.8%; Score 30; DB 10; Length 185;
Best Local Similarity 62.5%; Pred. No. 23;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NMLSEVER 8
DB 73 NMLKEIER 80

RESULT 5
US-09-815-242-5013
; Sequence 5013, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
```

```
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5013
; LENGTH: 584
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-815-242-5013

Query Match          69.8%; Score 30; DB 10; Length 584;
Best Local Similarity 66.7%; Pred. No. 85;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NMLSEVER 9
DB 255 NMLAEVME 263

RESULT 6
US-09-815-242-10878
; Sequence 10878, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10878
; LENGTH: 589
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-815-242-10878

Query Match          69.8%; Score 30; DB 10; Length 589;
Best Local Similarity 66.7%; Pred. No. 86;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NMLSEVER 9
DB 111:11:1
```

Db 255 NMLAEVWKE 263

RESULT 7

US-09-815-242-13724
; Sequence 13724, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13724
; LENGTH: 1294
; TYPE: PRT
; ORGANISM: Salmonella typhi
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(1294)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-815-242-13724

Query Match

Best Local Similarity 69.8%; Score 30; DB 10; Length 1294;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 NMLSEVERE 9

Db 1186 DVLAEIERE 1194

RESULT 8

US-09-263-811-9
; Sequence 9, Application US/09263811
; Patent No. US20020058610A1
; GENERAL INFORMATION:
; APPLICANT: Jian NI et al.
; TITLE OF INVENTION: Mammary Transforming Protein
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HUMAN GENOME SCIENCES, INC
; STREET: 9410 KEY WEST AVENUE
; CITY: ROCKVILLE
; STATE: MARYLAND
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE

COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/263,811
FILING DATE: MAR-08-1999
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/743,975
FILING DATE: 01-NOV-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/006,187
FILING DATE: 02 NOVEMBER 1995
ATTORNEY/AGENT INFORMATION:
NAME: MICHELE M. WALES
REGISTRATION NUMBER: 43,975
REFERENCE/DOCKET NUMBER: PF212D1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 301-309-8439
TELEFAX: 301-309-8504
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 61 AMINO ACIDS
TYPE: AMINO ACID
STRANDEDNESS:
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN

US-09-263-811-9

Query Match

Best Local Similarity 67.4%; Score 29; DB 10; Length 61;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 NMLSEVERE 9

Db 19 NLSRPERE 27

RESULT 9

US-09-250-611-2
; Sequence 2, Application US/09250611
; Patent No. US20020143161A1
; GENERAL INFORMATION:
; APPLICANT: Byrne, Jennifer A.
; APPLICANT: Bassett, Paul
; TITLE OF INVENTION: Members of the D52 Gene Family
; FILE REFERENCE: 1383.0210001
; CURRENT APPLICATION NUMBER: US/09/250,611
; CURRENT FILING DATE: 1999-02-17
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 144
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-250-611-2

Query Match

Best Local Similarity 67.4%; Score 29; DB 10; Length 144;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NMLSEVERE 9

Db 42 SMLSEERE 50

RESULT 10

US-09-250-611-4
; Sequence 4, Application US/09250611
; Patent No. US20020143161A1
; GENERAL INFORMATION:
; APPLICANT: Byrne, Jennifer A.
; APPLICANT: Bassett, Paul

;; TITLE OF INVENTION: Members of the D52 Gene Family
;; FILE REFERENCE: 1383.0210001
;; CURRENT APPLICATION NUMBER: US/09/250,611
;; CURRENT FILING DATE: 1999-02-17
;; NUMBER OF SEQ ID NOS: 108
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 4
;; LENGTH: 204
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-250-611-4

Query Match 67.4%; Score 29; DB 10; Length 204;
Best Local Similarity 66.7%; Pred. No. 42;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NMISEVERE 9
:|||||:|:
Db 29 SMLSEERE 37

RESULT 11
US-09-250-611-9
; Sequence 9, Application US/09250611
; Patent No. US20020143161A1
; GENERAL INFORMATION:
; APPLICANT: Byrne, Jennifer A.
; APPLICANT: Bassett, Paul
; TITLE OF INVENTION: Members of the D52 Gene Family
; FILE REFERENCE: 1383.0210001
; CURRENT APPLICATION NUMBER: US/09/250,611
; CURRENT FILING DATE: 1999-02-17
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 204
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-250-611-9

Query Match 67.4%; Score 29; DB 10; Length 204;
Best Local Similarity 66.7%; Pred. No. 42;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NMISEVERE 9
:|||||:|:
Db 29 SMLSEERE 37

RESULT 12
US-10-027-450-24
; Sequence 24, Application US/10027450
; Patent No. US20020102715A1
; GENERAL INFORMATION:
; APPLICANT: Falco, Saverio Carl
; APPLICANT: Hiltz, William D.
; APPLICANT: Kinney, Anthony J.
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Rafalski, J. Antoni
; TITLE OF INVENTION: PLANT BRANCHED CHAIN AMINO ACID BIOSYNTHETIC ENZYMES
; FILE REFERENCE: BB-1126
; CURRENT APPLICATION NUMBER: US/10/027,450
; CURRENT FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: 60/063,423
; PRIOR FILING DATE: 1997 October 28
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: Microsoft Word Version 7.0A
; SEQ ID NO 24
; LENGTH: 285
; TYPE: PRT
; ORGANISM: Glycine max
US-10-027-450-24

Query Match 67.4%; Score 29; DB 12; Length 285;
Best Local Similarity 66.7%; Pred. No. 61;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NMISEVERE 9
:|||||:|:
Db 142 NVLSKVEAE 150

RESULT 13
US-09-978-756-2
; Sequence 2, Application US/09978756
; Patent No. US20020160017A1
; GENERAL INFORMATION:
; APPLICANT: Holder, Anthony
; APPLICANT: Birdsell, Berry
; APPLICANT: Feeney, James
; APPLICANT: Morgan, William
; APPLICANT: Syed, Shabih
; TITLE OF INVENTION: Malaria Vaccine
; FILE REFERENCE: 18396/1005
; CURRENT APPLICATION NUMBER: US/09/978,756
; CURRENT FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: PCT/GB00/01558
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: 09/311,817
; PRIOR FILING DATE: 1999-05-13
; PRIOR APPLICATION NUMBER: 2,271,451
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 9909072.2
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 376
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
US-09-978-756-2

Query Match 67.4%; Score 29; DB 9; Length 376;
Best Local Similarity 55.6%; Pred. No. 83;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 NMISEVERE 9
:|||||:|:
Db 9 NILSKIENE 17

RESULT 14
US-09-895-913A-30
; Sequence 30, Application US/09895913A
; Patent No. US20020160456A1
; GENERAL INFORMATION:
; APPLICANT: Kieanthous, Harold
; APPLICANT: Al-Garawi, Amal
; APPLICANT: Miller, Charles
; APPLICANT: Tomb, Jean Francois
; APPLICANT: Oomen, Raymond P.
; TITLE OF INVENTION: Identification of Polynucleotides
; TITLE OF INVENTION: Encoding No. US20020160456A1a1 Helicobacter Polypeptides in t
; FILE REFERENCE: 06132/043002
; CURRENT APPLICATION NUMBER: US/09/895,913A
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 08/881,227
; PRIOR FILING DATE: 1997-06-24
; NUMBER OF SEQ ID NOS: 368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 419
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-09-895-913A-30

Query Match 67.4%; Score 29; DB 9; Length 419;
 Best Local Similarity 66.7%; Pred. No. 94;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NMLSEVERE 9
 DB 276 NSLSEVKOE 284

RESULT 15

US-09-895-913A-242
 ; Sequence 242, Application US/09895913A
 ; Patent No. US20020160456a1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kleanthous, Harold
 ; APPLICANT: Al-Garawi, Amal
 ; APPLICANT: Miller, Charles
 ; APPLICANT: Tomb, Jean Francois
 ; APPLICANT: Oomen, Raymond P.
 ; TITLE OF INVENTION: Identification of Polynucleotides
 ; TITLE OF INVENTION: Encoding No. US20020160456a1 Helicobacter Polypeptides in the
 ; TITLE OF INVENTION: Genome
 ; FILE REFERENCE: 06132/043002
 ; CURRENT APPLICATION NUMBER: US/09/895,913A
 ; CURRENT FILING DATE: 2001-06-29
 ; PRIOR APPLICATION NUMBER: US 08/881,227
 ; PRIOR FILING DATE: 1997-06-24
 ; NUMBER OF SEQ ID NOS: 368
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 242
 ; LENGTH: 419
 ; TYPE: PRT
 ; ORGANISM: Helicobacter pylori
 US-09-895-913A-242

Query Match 67.4%; Score 29; DB 9; Length 419;
 Best Local Similarity 66.7%; Pred. No. 94;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NMLSEVERE 9
 DB 276 NSLSEVKOE 284

Search completed: November 18, 2002, 16:05:44
 Job time : 4.93548 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 18, 2002, 15:53:18 ; Search time 10.7419 Seconds
(without alignments)
80.545 Million cell updates/sec

Title: US-09-848-781-1

Perfect score: 43

Sequence: 1 NMLSEVERE 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	38	88.4	510	2 A60280	bacillolysin-like
2	38	88.4	510	2 B60280	zinc metalloprotei
3	38	88.4	510	2 A01100	ryanodine receptor
4	38	88.4	5126	2 S40450	probable DNA mlsma
5	34	79.1	766	2 A75332	hypothetical prote
6	33	76.7	284	2 E83347	probable ABC trans
7	33	76.7	284	2 T45281	conserved hypothet
8	33	76.7	359	2 B69070	superfamily I DNA
9	33	76.7	721	2 B69074	myosin-like protei
10	33	76.7	784	2 D96564	preprotein transio
11	33	76.7	1036	2 A57386	centrin - human
12	32	74.4	172	2 I38424	hypothetical prote
13	32	74.4	388	2 T47530	hypothetical prote
14	32	74.4	409	2 S70119	hypothetical prote
15	32	74.4	636	2 T18777	hypothetical prote
16	32	74.4	645	2 T26926	hypothetical prote
17	32	74.4	1088	2 T18559	hypothetical prote
18	31	72.1	132	2 A00130	conserved hypothet
19	31	72.1	181	2 H64406	hypothetical prote
20	31	72.1	190	2 AG0030	conserved hypothet
21	31	72.1	190	2 G71672	hypothetical prote
22	31	72.1	200	2 T48976	hypothetical prote
23	31	72.1	277	2 G71021	hypothetical prote
24	31	72.1	333	2 G81132	NITF3/SMI family
25	31	72.1	341	2 T24927	hypothetical prote
26	31	72.1	353	2 G69436	hypothetical prote
27	31	72.1	366	2 F72062	hypothetical prote
28	31	72.1	366	2 A86561	CT449 hypothetical prote
29	31	72.1	372	2 H98072	hypothetical prote

30	31	72.1	429	2 H90157	aspartyl-tRNA synt
31	31	72.1	482	2 T02739	probable cytochrom
32	31	72.1	540	2 A71610	Hsp60 fold T-compl
33	31	72.1	594	1 A44073	CIR1 protein - yea
34	31	72.1	629	2 E64046	mismatch repair pr
35	31	72.1	758	2 E83884	hypothetical prote
36	30	69.8	53	2 A62231	hypothetical prote
37	30	69.8	151	2 E71547	hypothetical prote
38	30	69.8	184	2 T27050	hypothetical prote
39	30	69.8	185	1 H64676	ribosome releasing
40	30	69.8	185	2 F71837	ribosome releasing
41	30	69.8	187	2 T05071	hypothetical prote
42	30	69.8	209	2 T45109	H+-transporting tw
43	30	69.8	217	2 T27258	hypothetical prote
44	30	69.8	224	2 B56168	deoxyguanosine kin
45	30	69.8	238	2 T50846	hypothetical prote

ALIGNMENTS

RESULT 1

A60280
Bacillolysin-like proteinase (EC 3.4.24.-) prtA precursor - *Listeria monocytogenes* (s

N:Alternate names: microbial metalloproteinase mp1

C:Species: *Listeria monocytogenes*

A:Variety: strain EGD, serotype 1/2a

C:Date: 03-Mar-1993 #sequence-revision 03-Mar-1993 #text-change 22-Jun-1999

C:Accession: A43575; C60280; A60280; B37204

R:Domann, E.; Jelmisteer-Waechter, M.; Goebel, W.; Chakraborty, T.

A:Title: Molecular cloning, sequencing, and identification of a metalloproteinase gene

A:Reference number: A43575; MUID:91100010; PMID:1898903

A:Accession: A43575

A:Molecule type: DNA

A:Residues: 1-510 <DOM>

A:Cross-References: GB:X54619; NID:944114; PIDN:CA48439.1; PID:944115

A:Experimental source: strain EGD, serotype 1/2a

R:Mengaud, J.; Geoffroy, C.; Cossart, P.

Infect. Immun. 59, 1043-1049, 1991

A:Title: Identification of a new operon involved in *Listeria monocytogenes* virulence:

A:Reference number: A60280; MUID:91147180; PMID:1705239

A:Accession: C60280

A:Status: not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-185, 'A', 187-426, 'I', 428-510 <ME2>

A:Experimental source: strain EGD, serotype 1/2a

A:Accession: A60280

A:Status: not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-510 <MEN>

A:Experimental source: strain 1028, serotype 1/2c

R:Mengaud, J.; Vicente, M.F.; Cossart, P.

Infect. Immun. 57, 3695-3701, 1989

A:Title: Transcriptional mapping and nucleotide sequence of the *Listeria monocytogene*

A:Reference number: A37204; MUID:90035452; PMID:2509367

A:Accession: B37204

A:Molecule type: DNA

A:Residues: 1-47 <ME3>

C:Genetics:

A:Gene: prtA, mp1

C:Superfamily: thermolysin

C:Keywords: extracellular protein; hydrolase; metalloproteinase; zinc

F:1-24/Domain: signal sequence #status predicted <SIG>

F:25-510/Product: bacillolysin-like proteinase #status predicted <MAT>

Query Match 88.4%; Score 38; DB 2; Length 510;

Best Local Similarity 100.0%; Pred. No. 7.5;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NMLSEVER 8

DB 196 NMLSEVER 203

RESULT 2

B60280

bacillolysin-like proteinase (EC 3.4.24.-) prta precursor - *Listeria monocytogenes* (strain)

N:Alternate names: metalloproteinase homolog mpr

C:Species: *Listeria monocytogenes*

C>Date: 03-Mar-1993 #sequence_revision 03-Mar-1993 #text_change 15-Oct-1999

C:Accession: B60280; A43868; S24232

R:Mengaud, J.; Geoffroy, C.; Cosart, P.

Infect. Immun. 59, 1043-1049, 1991

A:Title: Identification of a new operon involved in *Listeria monocytogenes* virulence: it

A:Reference number: A60280; MUID:91147180; PMID:1705239

A:Accession: B60280

A>Status: not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-510 <MEN>

A:Experimental source: strain LM8, serotype 4b

R:Vazquez-Boland, J.A.; Kocks, C.; Dramsl, S.; Ohayon, H.; Geoffroy, C.; Mengaud, J.; Co

Infect. Immun. 60, 219-230, 1992

A:Title: Nucleotide sequence of the lecithinase operon of *Listeria monocytogenes* and pos

A:Reference number: A43868; MUID:92104678; PMID:1309513

A:Accession: A43868

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 504-510 <VAZ>

A>Note: sequence extracted from NCBI backbone (NCBI:74437, NCBI:74457)

R:Rasmussen, O.F.; Beck, T.; Olsen, J.E.; Dons, L.; Rossen, L.

Infect. Immun. 59, 3945-3951, 1991

A:Title: *Listeria monocytogenes* isolates can be classified into two major types accordit

A:Reference number: S24230; MUID:92040062; PMID:1937753

A:Accession: S24232

A>Status: nucleotide sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-46, 'A', 48-102, 'A', 104-271 <RAS>

A:Cross-references: EMBL:X60035

A:Experimental source: strain 12067, serotype4b

A>Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1991

C:Genetics:

A:Gene: prta; mpl

C:Superfamily: thermolysin

C:Keywords: extracellular protein; hydrolase; metalloproteinase; zinc

F:1-24/Domains: signal sequence #status predicted <SIG>

F:5-510/Product: bacillolysin-like proteinase #status predicted <MAT>

Query Match 88.4%; Score 38; DB 2; Length 510;

Best Local Similarity 100.0%; Pred. No. 7.5;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NMLSEVER 8

Db 196 NMLSEVER 203

RESULT 3

AD1100

zinc metalloproteinase precursor [imported] - *Listeria monocytogenes* (strain EGD-e)

C:Species: *Listeria monocytogenes*

C>Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001

C:Accession: AD1100

R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker

; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entlian, K.D.; Fsihl, H.

D.; Jones, L.M.; Karst, U.

Science 294, 849-852, 2001

A:Authors: Kref, J.; Kuhn, M.; Kunst, F.; Kurapatk, G.; Madueno, E.; Maltournam, A.; Ma

Ok, C.; Schluter, T.; Simoes, N.; Tiller, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlund,

A:Title: Comparative genomics of *Listeria* species.

A:Reference number: AB1077; MUID:21537279; PMID:11679669

A:Accession: AD1100

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-510 <GLA>

A:Cross-references: GB:NC_003210; PIDN:CAD00730.1; PID:916409568; GSPDB:GN00177

A:Experimental source: strain EGD-e

C:Genetics:

A:Gene: mpl

C:Superfamily: thermolysin

Query Match 88.4%; Score 38; DB 2; Length 510;

Best Local Similarity 100.0%; Pred. No. 7.5;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NMLSEVER 8

Db 196 NMLSEVER 203

RESULT 4

S40450

ryanodine receptor/calcium release channel - fruit fly (*Drosophila melanogaster*)

C:Species: *Drosophila melanogaster*

C>Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 11-Jun-1999

C:Accession: S40450

R:Takeshima, H.; Nishi, M.; Iwabe, N.; Miyata, T.; Hosoya, T.; Masai, I.; Hotta, Y.

FEBS Lett. 337, 81-87, 1994

A:Title: Isolation and characterization of a gene for a ryanodine receptor/calcium re

A:Reference number: S40450; MUID:94102409; PMID:8276118

A:Accession: S40450

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-5126 <TAK>

C:Superfamily: ryanodine receptor; transcription initiation factor sigma region 1 hom

Query Match 88.4%; Score 38; DB 2; Length 5126;

Best Local Similarity 88.9%; Pred. No. 89;

Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NMLSEVER 9

Db 2890 NMLSEVER 2898

RESULT 5

A75332

probable DNA mismatch repair protein Muts - *Deinococcus radiodurans* (strain R1)

C:Species: *Deinococcus radiodurans*

C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000

C:Accession: A75332

R:White, O.; Eisen, J.A.; Heldelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.

; M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.

Science 286, 1571-1577, 1999

A:Title: Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1.

A:Reference number: A75250; MUID:20036896; PMID:10567266

A:Accession: A75332

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-766 <WHI>

A:Cross-references: GB:AE002035; GB:AE00513; NID:96459753; PIDN:AAF11527.1; PID:9645

A:Experimental source: strain R1

C:Genetics:

A:Gene: DR1976

A:Map position: 1

Query Match 79.1%; Score 34; DB 2; Length 766;

Best Local Similarity 87.5%; Pred. No. 77;

Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 MUSEVER 9

Db 513 MUSEVER 520

RESULT 6

E83347

hypothetical protein PA2395 [imported] - *Pseudomonas aeruginosa* (strain PA01)

C:Species: *Pseudomonas aeruginosa*
C:Date: 15-Sep-2000 #sequence_rev: 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: E83347
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; Lam,
; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic patho
A:Reference number: AB2950; MUID:20437337; PMID:10984043
A:Accession: E83347
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-284 <STO>
A:Cross-references: GB:AE004665; GB:AE004091; NID:99948426; PIDN:AAG05783.1; GSPDB:GN001
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA2395

Query Match 76.7%; Score 33; DB 2; Length 284;
Best Local Similarity 87.5%; Pred. No. 42;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 MISEVERE 9
|||||
Db 149 MISEVERE 156

RESULT 7
T45281
probable ABC transporter intracellular ATPase chain bldKE [imported] - Streptomyces coel
C:Species: Streptomyces coelicolor
C:Date: 31-Jan-2000 #sequence_rev: 31-Jan-2000 #text_change 31-Jan-2000
C:Accession: T45281
R:Nozwell J., R.; McGovern, K.; Losick, R.
submitted to the EMBL Data Library, August 1996
A:Description: An oligopeptide permease responsible for the import of an extracellular s
A:Reference number: Z22954
A:Accession: T45281
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-284 <NMD>
A:Cross-references: EMBL:U68036; PIDN:AAB09558.1
A:Experimental source: strain M145
C:Genetics:
A:Gene: bldKE
C:Function:
A:Description: involved in the formation of an aerial mycelium
C:Keywords: oligopeptide transport

Query Match 76.7%; Score 33; DB 2; Length 284;
Best Local Similarity 66.7%; Pred. No. 42;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MISEVERE 9
|:|:|:|
Db 208 MISEVERE 216

RESULT 8
B69070
conserved hypothetical protein MTH1522 - Methanobacterium thermoautotrophicum (strain De
C:Species: Methanobacterium thermoautotrophicum
C:Date: 05-Dec-1997 #sequence_rev: 05-Dec-1997 #text_change 22-Oct-1999
C:Accession: B69070
R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.
K.; S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct
A:Reference number: A69000; MUID:98037514; PMID:9371463
A:Accession: B69070
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA

```

A:Residues: 1-159 <NMH>
A:Cross-references: GB:AE000912; GB:AE000666; NID:g2622632; PID:AA85997.1; PID:g262
A:Experimental source: strain Delta H
C:Genetics:
A:Gene: MTH1522
A:Start codon: TTG

Query Match          76.7%; Score 33; DB 2; Length 359;
Best Local Similarity 75.0%; Pred. NO. 55;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY      1 NMLSEVER 8
          1:111:11
Db      246 NILSEIER 253

RESULT 9
B96974
superfamily I DNA and RNA helicase [imported] - Clostridium acetobutylicum
C:Species: Clostridium acetobutylicum
C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 30-Sep-2001
C:Accession: B96974
R:Rolling, J.; Bennett, G.; Omeichenko, M.V.; Markarova, K.S.; zeng, Q.; Gibson, R.; L
J. Bacteriol. 183, 4623-4838, 2001
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium
A:Reference number: A96900; MUID:21359325; PMID:21359325
A:Accession: B96974
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-721 <KUR>
A:Cross-references: GB:AE001437; PIDN:AAK78581.1; PID:g15023473; GSPDB:GN00168
A:Experimental source: Clostridium acetobutylicum ATCC824
C:Genetics:
A:Gene: CAC0603
C:Superfamily: Bacillus subtilis hypothetical protein yvgs

Query Match          76.7%; Score 33; DB 2; Length 721;
Best Local Similarity 55.6%; Pred. No. 1.2e+02;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY      1 NMLSEVER 9
          1:111:11
Db      349 NLVSEIERO 357

RESULT 10
D96564
myosin-like protein, 12311-7712 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: D96564
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon
Chiu, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,
ansen, N.F.; Hughes, B.; Huizart, L.
Nature 408, 816-820, 2000
A:Author: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzia
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: D96564
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-784 <STO>
A:Cross-references: GB:AE005173; NID:g10645443; PIDN:AA621559.1; GSPDB:GN00141
C:Genetics:
A:Gene: F19K6.14
A:Map position: 1

Query Match          76.7%; Score 33; DB 2; Length 784;

```

Best Local Similarity 66.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 NMLSEVERE 9
Db 438 NMLSEVERE 446

RESULT 11

A57386
preprotein translocase secA precursor - spinach

C:Species: Spinacia oleracea (spinach)
C>Date: 08-Dec-1995 #sequence_revision 08-Dec-1995 #text_change 17-Nov-2000

C:Accession: A57386

R:Berghoefer, J.; Kanauchov, I.; Herrmann, R.G.; Kloesgen, R.B.

J. Biol. Chem. 270, 18341-18346, 1995

A:Title: Isolation and characterization of a cDNA encoding the SecA protein from spinach

A:Reference number: A57386; MUID:95355455; PMID:7629156

A:Accession: A57386

A:Status: preliminary; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-1036 <BER>

A:Cross-references: GB:249124

C:Comment: The "nucleotide-binding motif B" and "DEAD motif" features as annotated are adjacent to the identified motif and a third conserved motif is approximately 120-140

C:Superfamily: preprotein translocase secA

C:Keywords: ATP; chloroplast; membrane-associated complex; P-loop; protein transport

F:1-84/Domain: trans peptide (chloroplast) #status predicted <TN>

F:85-1036/Product: preprotein translocase secA #status predicted <MAT>

F:186-193/Region: nucleotide-binding motif A (P-loop) #status atypical

F:289-294/Region: nucleotide-binding motif B

F:293-296/Region: DEAD motif

Query Match

Best Local Similarity 76.7%; Score 33; DB 2; Length 1036;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 NMLSEVERE 9
Db 244 NMLSEVERE 252

RESULT 12

I38424
centrin - human

C:Species: Homo sapiens (man)

C>Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 21-Jul-2000

C:Accession: I38424

R:Erradiou, R.; Sanders, M.A.; Salisbury, J.L.

J. Cell Sci. 107, 9-16, 1994

A:Title: Cloning of a cDNA encoding human centrin, an EF-hand protein of centrosomes and

A:Reference number: I38424; MUID:94220620; PMID:8175926

A:Accession: I38424

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-172 <RES>

A:Cross-references: EMBL:U03270; NID:9333734; PIDN:AAC27343.1; PID:9414993

C:Superfamily: calmodulin; calmodulin repeat homology

C:Keywords: EF hand

F:28-60/Domain: calmodulin repeat homology <EF1>

F:64-96/Domain: calmodulin repeat homology <EF2>

F:101-133/Domain: calmodulin repeat homology <EF3>

F:137-169/Domain: calmodulin repeat homology <EF4>

Query Match
Best Local Similarity 74.4%; Score 32; DB 2; Length 172;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 2 NMLSEVERE 9
Db 72 NMLSEVERE 79

RESULT 13
T47530
hypothetical protein F16L2.110 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000

C:Accession: T47530

R:Jordan, N.; Bangert, S.; Wiedelmann, R.; Voss, H.; Unseld, M.; Mewes, H.W.; Rudd, S

submitted to the Protein Sequence Database, March 2000

A:Reference number: 224468

A:Accession: T47530

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-388 <TOR>

A:Cross-references: EMBL:AL162459

A:Experimental source: cultivar Columbia; BAC clone F16L2

C:Genetics:

A:Map position: 3

A:Introns: 189/3; 208/2; 272/3

A>Note: F16L2.110

Query Match
Best Local Similarity 74.4%; Score 32; DB 2; Length 388;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 NMLSEVERE 9
Db 125 NMLSEVERE 133

RESULT 14

S70119
hypothetical protein YDR289c - yeast (Saccharomyces cerevisiae)

N:Alternate names: hypothetical protein D9819.2

C:Species: Saccharomyces cerevisiae

C>Date: 24-Aug-1996 #sequence_revision 06-Sep-1996 #text_change 19-Apr-2002

C:Accession: S70119

R:Fulton, L.

submitted to the EMBL Data Library, May 1996

A:Description: The sequence of S. cerevisiae cosmid 9819.

A:Reference number: S70114

A:Accession: S70119

A:Molecule type: DNA

A:Residues: 1-409 <FUL>

A:Cross-references: EMBL:U51031; NID:91332635; PID:91230654; GSPDB:GN00004; MIPS:YDR2

C:Genetics:

A:Gene: SGD:RTT103; MIPS:YDR289C

A:Cross-references: SGD:S0002697

A:Map position: 4R

Query Match
Best Local Similarity 74.4%; Score 32; DB 2; Length 409;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 NMLSEVERE 7
Db 233 NMLSEVERE 239

RESULT 15

T18777
hypothetical protein B0513.6 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 01-Mar-2002

C:Accession: T18777

R:Matthews, L.

submitted to the EMBL Data Library, November 1996

A:Reference number: Z19020

A:Accession: T18777

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-636 <WIL>

A:Cross-references: EMBL:282256; PIDN:CAB05118.1; GSPDB:GN00022; CESP:B0513.6

A: Experimental source: clone B0513

C: Genetics:

A: Gene: CESP.B0513.6

A: Map position: 4

A: Introns: 50/2; 88/2; 171/2; 209/2; 290/1; 324/2; 543/3; 614/3

C: Superfamily: Caenorhabditis elegans hypothetical protein Y57A10A.d

Query Match 74.4%; Score 32; DB 2; Length 636;

Best Local Similarity 66.7%; Pred. No. 1.6e+02;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NMLSEVERE 9

Db 320 NMLSEVERE 328

Search completed: November 18, 2002, 15:56:07

Job time : 12.7419 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compen Ltd.

OM protein - protein search, using sw model

Run on: November 18, 2002, 15:53:17 ; Search time 5.80645 Seconds
(without alignments)
64.288 Million cell updates/sec

Title: US-09-848-781-1
Perfect score: 43
Sequence: 1 NMLSEVERE 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	38	88.4	510	1	PRO1_LISMO
2	38	88.4	510	1	PRO2_LISMO
3	33	76.7	359	1	YF22_METTH
4	33	76.7	1036	1	SECA_SPTOL
5	32	74.4	172	1	CAT2_HUMAN
6	32	74.4	896	1	Y249_HUMAN
7	31	72.1	151	1	PFDA_SULTO
8	31	72.1	181	1	Y856_METUA
9	31	72.1	594	1	CIR1_YEAST
10	31	72.1	629	1	MUTL_HAEIN
11	30	69.8	185	1	RRF_HELPJ
12	30	69.8	185	1	RRF_HELPJ
13	30	69.8	209	1	VARD_METMA
14	30	69.8	223	1	DGR2_LACAC
15	30	69.8	378	1	PTT6_BOVIN
16	30	69.8	390	1	DXR_FUSNN
17	30	69.8	840	1	UREA_CANEN
18	30	69.8	4451	1	GRSB_BACBR
19	29	67.4	95	1	RPOL_PYRHO
20	29	67.4	96	1	C553_HELPY
21	29	67.4	121	1	RIL9_CHLPN
22	29	67.4	148	1	CAUL_HUMAN
23	29	67.4	172	1	CATR_MOUSE
24	29	67.4	180	1	ATPF_ORYSA
25	29	67.4	183	1	ATPF_MAIZE
26	29	67.4	183	1	ATPF_WHEAT
27	29	67.4	184	1	ATPF_ARATH
28	29	67.4	184	1	ATPF_SPTOL
29	29	67.4	184	1	ATPF_TOBAC
30	29	67.4	192	1	DYR_CANAL
31	29	67.4	204	1	TD53_HUMAN
32	29	67.4	204	1	TD53_MOUSE
33	29	67.4	204	1	TRPF_SULSO

34	29	67.4	240	1	HEM4_METUA	Q58401 methanococ
35	29	67.4	247	1	YF01_ARCFU	O28771 archaeoglob
36	29	67.4	278	1	EXSA_PSEAE	P26593 pseudomonas
37	29	67.4	292	1	BLAN_ENTCL	P52653 enterobacte
38	29	67.4	379	1	PANE_YEAST	P38787 saccharomy
39	29	67.4	403	1	Y271_AOUAE	O66628 aquilex aeo
40	29	67.4	410	1	VCLB_PEA	P02854 plasm sativ
41	29	67.4	466	1	SYC_CLOPE	O88955 clostridium
42	29	67.4	485	1	CG2A_CAEEL	P34638 caenorhabd1
43	29	67.4	744	1	CATA_EMENT	P55305 emerocella
44	29	67.4	802	1	MUTS_BUCAI	P57504 buchneera ap
45	29	67.4	853	1	MUTS_AZOV1	P27345 azotobacter

ALIGNMENTS

RESULT 1
PRO1_LISMO STANDARD; PRT; 510 AA.
AC P23224;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Zinc metalloproteinase precursor (EC 3.4.24.-).
GN MPL OR PRTA OR LMO0203.
OS Listeria monocytogenes.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_Taxid=1639;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-EGD / Serovar 1/2a;
RX MEDLINE=91100010; PubMed=1898903;
RA Donnann E., Leimeister-Waecheher M., Goebel W., Chakraborty T.;
RT "Molecular cloning, sequencing, and identification of a
metallopeptidase gene from Listeria monocytogenes that is species
specific and physically linked to the listeriolysin gene.";
RL Infect. Immun. 59:65-72(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-EGD-e / Serovar 1/2a;
RX MEDLINE=21537279; PubMed=11679669;
RA Glaser P., Franjeul L., Buchrieser C., Rusnlok C., Amend A.,
BA Baguerio F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA Donnann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussauget O.,
RA Entian K.-D., Eschl H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Kretz J., Kunz M., Kunst F., Kurapkhat G.,
RA Madueno E., Maitournem A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordstedt G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
RT "Comparative genomics of Listeria species.";
RL Science 294:849-852(2001).
CC -I- FUNCTION: PROBABLY LINKED TO THE PATHOGENESIS OF LISTERIAL
INFECTION.
CC -I- COFACTOR: BINDS 1 ZINC ION.
CC -I- SUBCELLULAR LOCATION: Secreted.
CC -I- INDUCTION: THE MPL AND THE LISTERIOLYSIN GENES BEING PHYSICALLY
LINKED, THEIR EXPRESSION MAY BE REGULATED IN A SIMILAR MANNER.
CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M4.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
DR EMBL: Y54619; CAA38439.1; -;
DR EMBL: AL591974; CAD00730.1; -;

```
DR PIR; A43575; A43575.
DR PIR; A60280; A60280.
DR HSSP; P00800; 1HYR.
DR MEROPS; M04.008; -.
DR ListList; LMO00203; -.
DR InterPro; IPR005075; Pep_M4-propep.
DR InterPro; IPR001570; Peptidase_M4.
DR InterPro; IPR000130; Zn_MTPeptidse.
DR Pfam; PF01447; Peptidase_M4; 1.
DR Pfam; PF02868; Peptidase_M4.C; 1.
DR Pfam; PF03413; Pep_M4-propep; 1.
DR PRINTS; PR00730; THERMOLYSIN.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
DR HydroLase; Metalloprotease; Zinc; Zymogen; signal; Complete proteome.
KW SIGNAL 1 24
FT PROPEP 25 200
FT CHAIN 201 510
FT METAL 349 349
FT ACT_SITE 350 350
FT METAL 353 353
FT METAL 373 373
FT ACT_SITE 437 437
SQ SEQUENCE 510 AA; 57411 MW; FF978FBDCA804C0 CRC64;
```

```
Query Match
Best Local Similarity 88.4%; Score 38; DB 1; Length 510;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 NMLSEVER 8
DB 196 NMLSEVER 203
```

```
RESULT 2
PROJ.LISMO STANDARD; PRT; 510 AA.
ID PROJ.LISMO
AC P34025;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Zinc metalloproteinase precursor (EC 3.4.24.-).
GN MPL OR PRTA.
OS Listeria monocytogenes.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
ON NCBI_TaxID=1639;
RX SEQUENCE FROM N.A.
RC STRAIN=LO28 / Serovar 1/2c;
RA MEDLINE=91147180, PubMed=1705239;
RA Mengaud J., Geoffroy C., Cossart P.;
RT Identification of a new operon involved in Listeria monocytogenes
RT virulence: its first gene encodes a protein homologous to bacterial
RT metalloproteases."
RL Infect. Immun. 59:1043-1049(1991).
RN [2]
RP SEQUENCE OF 1-272 FROM N.A.
RC STRAIN=12067;
RX MEDLINE=92040062; PubMed=1937753;
RA Rasmussen O.F., Beck T., Olsen J.E., Dons L., Rossen L.;
RT Listeria monocytogenes isolates can be classified into two major
RT types according to the sequence of the listeriolysin gene."
RL Infect. Immun. 59:3945-3951(1991).
CC -1- FUNCTION: PROBABLY LINKED TO THE PATHOGENESIS OF LISTERIAL
CC INFECTION.
CC -1- COFACTOR: BINDS 1 ZINC ION.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- INDUCTION: THE MPL AND THE LISTERIOLYSIN GENES BEING PHYSICALLY
CC LINKED, THEIR EXPRESSION MAY BE REGULATED IN A SIMILAR MANNER.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M4.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
```

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

```
CC EMBL; X60035; CAA42640.1; -.
DR PIR; B60280; B60280.
DR PIR; S24232; S24232.
DR HSSP; P00800; 1HYR.
DR MEROPS; M04.008; -.
DR InterPro; IPR005075; Pep_M4-propep.
DR InterPro; IPR001570; Peptidase_M4.
DR InterPro; IPR000130; Zn_MTPeptidse.
DR Pfam; PF01447; Peptidase_M4; 1.
DR Pfam; PF02868; Peptidase_M4.C; 1.
DR Pfam; PF03413; Pep_M4-propep; 1.
DR PRINTS; PR00730; THERMOLYSIN.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
DR HydroLase; Metalloprotease; Zinc; Zymogen; signal.
KW SIGNAL 1 24
FT PROPEP 25 200
FT CHAIN 201 510
FT METAL 349 349
FT ACT_SITE 350 350
FT METAL 353 353
FT METAL 373 373
FT ACT_SITE 437 437
FT CONFLICT 47 47
FT CONFLICT 103 103
SQ SEQUENCE 510 AA; 57569 MW; C166CB56515B175 CRC64;
```

```
Query Match
Best Local Similarity 88.4%; Score 38; DB 1; Length 510;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 NMLSEVER 8
DB 196 NMLSEVER 203
```

```
RESULT 3
YF22.METHH STANDARD; PRT; 359 AA.
ID YF22.METHH
AC O27566;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MTH1522.
GN MTH1522.
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
OC Methanobacteriaceae; Methanothermobacter.
ON NCBI_TaxID=187420;
RX SEQUENCE FROM N.A.
RC STRAIN=Delta H;
RX MEDLINE=98037514; PubMed=9371463;
RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
RA Harrison D., Hoang L., Keagle P., Lum W., Pothier B., Qiu D.,
RA Spadafora R., Vicare R., Wang Y., Wlezbowski J., Gibson R.,
RA Jiwani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
RA McDougall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,
RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
RT *Complete genome sequence of Methanobacterium thermoautotrophicum
RT deltaH: functional analysis and comparative genomics."
RL J. Bacteriol. 179:7135-7155(1997).
CC -1- SIMILARITY: STRONG, TO M.JANNSCHIT MJ1423.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE000912; AAB85997.1; -
DR InterPro: IPR000510; Oxred_nitrogsel.
DR Pfam: PF00148; oxidored_nitro.1.
RW Hypothetical protein; Complete proteome.
SQ SEQUENCE 359 AA; 39101 MW; 12B1FF882D80D28F CRC64;

Query March 76.7%; Score 33; DB 1; Length 359;
Best Local Similarity 75.0%; Pred. No. 22;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 NMLSEVER 8
Db 246 NMLSEIER 253
|:|:|:|:|
|:|:|:|:|

RESULT 4
SECA_SPTOL STANDARD; PRT: 1036 AA.
ID SECA_SPTOL STANDARD; PRT: 1036 AA.
AC 036795;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Preprotein translocase seca subunit, chloroplast precursor.
OS SECA.
CC Spinacia oleracea (Spinach).
CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
CC Caryophyllales; Caryophyllaceae; Spinacia.
OX NCBI_TaxID=3562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Monatel; TISSUE=leaf;
RX MEDLINE=95355455; PubMed=7629156;
RA Beeghoffer J., Karauchoy I., Herrmann R.G., Kloesgen R.B.;
RT "Isolation and characterization of a cDNA encoding the SecA protein
RT from spinach chloroplasts. Evidence for azide resistance of
RT Sec-dependent protein translocation across thylakoid membranes in
RT spinach".
RL J. Biol. Chem. 270:18341-18346(1995).
CC -1- FUNCTION: INVOLVED IN PROTEIN EXPORT. LIKELY PARTICIPATES IN
CC PROTEIN TRANSLOCATION ACROSS THE THYLAKOID MEMBRANE IN
CC CHLOROPLAST.
CC -1- SUBCELLULAR LOCATION: CHLOROPLAST STROMA. A MINOR FRACTION IS
CC ASSOCIATED WITH THE THYLAKOID MEMBRANE.
CC -1- SIMILARITY: BELONGS TO THE SECA FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z49124; CAAB8933.1; -
DR InterPro: IPR000185; SecA.
DR Pfam: PF01043; SecA_protein; 1.
DR PRINTS: PR00906; SECA.
DR TIGRfams: TIGR00963; seca; 1.
DR PROSITE: PS00312; SECA; 1.
RW Protein transport; ATP-binding; Chloroplast; Translocation; Transport;
KW Transit peptide.
FT TRANSIT 1 776 CHLOROPLAST (POTENTIAL).
FT CHAIN 777 1036 PREPROTEIN TRANSLOCASE SECA SUBUNIT.
FT NP_BIND 186 193 ATP (POTENTIAL).
SQ SEQUENCE 1036 AA; 116608 MW; 23920878B49A3283 CRC64;

Query Match 76.7%; Score 33; DB 1; Length 1036;
Best Local Similarity 77.8%; Pred. No. 65;

Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 NMLSEVER 9
Db 244 NMLSEVERE 252
|:|:|:|:|
|:|:|:|:|

RESULT 5
CAT2_HUMAN STANDARD; PRT: 172 AA.
ID CAT2_HUMAN STANDARD; PRT: 172 AA.
AC 012798;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Caltractin, isoform 2 (Centrin).
DE Caltractin, isoform 2 (Centrin).
GN CETN1 OR CETN OR CEN1.
OS Homo sapiens (Human).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=94230620; PubMed=8175926;
RA Errabolu R., Sanders M.A., Salisbury J.L.;
RT "Cloning of a cDNA encoding human centrin, an EF-hand protein of
RT centrosomes and mitotic spindle poles".
RL J. Cell Sci. 107:9-16(1994).
CC -1- FUNCTION: PLAYS A FUNDAMENTAL ROLE IN MICROTUBULE-ORGANIZING
CC CENTER STRUCTURE AND FUNCTION.
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: CENTROSOME OF INTERPHASE AND MITOTIC CELLS.
CC -1- MISCELLANEOUS: BINDS TWO MOLES OF CALCIUM PER MOLE OF PROTEIN.
CC -1- SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS, BELONGS
CC TO THE CALTRACTIN/CDC31 SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U03270; AAC27343.1; -
DR HSSP: P02593; ICDM.
DR Genew: HGNC:1866; CETN1.
DR MIM: 603187; -
DR InterPro: IPR002048; EF-hand.
DR Pfam: PF00036; efhand; 4.
DR PRODOM: PD000012; EF-hand; 2.
DR SMART: SM00054; EFh; 4.
DR PROSITE: PS00018; EF_HAND; 2.
KW Calcium-binding; Repeat; Cell division; Mitosis.
FT CA_BIND 41 52 EF_HAND 1 (PROBABLE).
FT DOMAIN 77 88 ANCESTRAL CALCIUM SITE 2 (POTENTIAL).
FT DOMAIN 114 125 ANCESTRAL CALCIUM SITE 3 (POTENTIAL).
FT CA_BIND 150 161 EF_HAND 4 (PROBABLE).
SQ SEQUENCE 172 AA; 19570 MW; 9ED8193F26F18881 CRC64;

Query Match 74.4%; Score 32; DB 1; Length 172;
Best Local Similarity 75.0%; Pred. No. 16;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 MISEVERE 9
Db 72 MISEVDRE 79
|:|:|:|:|
|:|:|:|:|

RESULT 6
Y249_HUMAN STANDARD; PRT: 896 AA.
ID Y249_HUMAN STANDARD; PRT: 896 AA.
AC Q92539;

```
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein KIAA0249.
GN KIAA0249.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
ON NCBI_TaxID=9606;
RX MEDLINE=97191544; PubMed=9039502;
RA Nagase T., Seki N., Ishikawa K.-I., Ohira M., Kawarabayashi Y.,
RA Ohara O., Tanaka A., Kotani H., Miyajima N., Nomura N.;
RT "Prediction of the coding sequences of unidentified human genes. VI.
RT The coding sequences of 80 new genes (KIAA0201-KIAA0280) deduced by
RT analysis of cDNA clones from cell line KG-1 and brain."
RL DNA Res. 3:321-328(1996).
CC -1- SIMILARITY: TO HUMAN KIAA0188 AND YEAST SMP2.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D87436; BAA13380.1; -
DR Genbank: HGNC:14450; LPTN2.
KW Hypothetical protein.
SQ SEQUENCE 896 AA; 99399 MW; 080113FCA533272 CRC64;

Query Match
Best Local Similarity 74.4%; Score 32; DB 1; Length 896;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 1 NMLSEVERE 9
DB 306 NMLSEVERE 314

RESULT 7
PDBA_SULTO STANDARD: PRT; 151 AA.
AC 097116;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Prefoldin alpha subunit (GIMC alpha subunit).
GN PDBA OR S71371.
OS Sulfolobus tokodaii.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
ON NCBI_TaxID=111955;
RX MEDLINE=21456156; PubMed=11572479;
RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
RA Sekine M., Baba S.-I., Ankei A., Kosugi H., Hosoyama A., Fukui S.,
RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
RA Yoshizawa T., Tanaka T., Yamazaki J., Kushida N., Oguchi A.,
RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
RA Oshima T., Kikuchi H.;
RT "Complete genome sequence of an aerobic thermophilic
RT Crenarchaeon, Sulfolobus tokodaii strain7."
RL DNA Res. 8:123-140(2001).
CC -1- FUNCTION: Molecular chaperone capable of stabilizing a range of
CC proteins. Seems to fulfil an ATP-independent, HSP70-like function
CC in archaeal de novo protein folding (By similarity).
CC -1- SUBUNIT: Heterohexamer of two alpha and four beta subunits (By
```

```
CC similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE PREFOLDIN ALPHA SUBUNIT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AP000986; BAB66434.1; ALT_INIT.
DR InterPro: IPR004127; DUF232.
DR Pfam: PF02996; DUF232; 1.
KW Chaperone; Complete proteome.
SQ SEQUENCE 151 AA; 16959 MW; A7B89A1C7659ED05 CRC64;
```

```
Query Match
Best Local Similarity 72.1%; Score 31; DB 1; Length 151;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 NMLSEVERE 9
DB 113 NMLSEVERE 121
```

```
RESULT 8
PDBA_METJA STANDARD: PRT; 181 AA.
AC 058266;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MJ0856.
GN MJ0856.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
ON NCBI_TaxID=2190;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Funtman J.L., Nguyen D.,
RA Uitterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Moese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii."
RL Science 273:1058-1073(1996).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U67529; AAB98860.1; -.
DR TIGR: MJ0856; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 181 AA; 21124 MW; DBB6B5BCD8F9F56 CRC64;
```

```
Query Match
Best Local Similarity 72.1%; Score 31; DB 1; Length 181;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
```

QY 1 NMLSEVERE 9
1:1:1111
Db 61 NMLSEVERE 69

RESULT 9
CIR1_YEAST
ID CIR1_YEAST STANDARD: PRT; 594 AA.
AC 001649;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Spindle pole body associated protein.
GN CIR1 OR YMR198W OR YMR646.11.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;

SEQUENCE FROM N.A.
[1]
RX MEDLINE=92354905; PubMed=1644287;
RA Page B.D., Snyder M.;
RT "CIR1: a developmentally regulated spindle pole body-associated
RT protein important for microtubule functions in Saccharomyces
RT cerevisiae.";
RL Genes Dev. 6:1414-1429(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / AB972;
RA Pearson D., Bowman S., Bartell B.G., Rajandream M.A.;
Submitted (JAN1995) to the EMBL/GenBank/DBJ databases.
CC -I- FUNCTION: A DEVELOPMENTALLY REGULATED PROTEIN IMPORTANT FOR
CC MICROTUBULE FUNCTIONS. TIGHTLY ASSOCIATED WITH KAR3, MAY SERVE TO
CC REGULATE THE CELLULAR COMPARTMENT IN WHICH KAR3 FUNCTIONS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC
CC -----
DR EMBL: M96439; AAA34494.1; -;
DR EMBL: Z47815; CAA87820.1; -;
DR PIR: A44073; A44073.
DR SGD: S0004811; CIR1.
KW Microtubules.
SO SEQUENCE 594 AA; 69069 MW; A2A09DEB8C2EC838 CRC64;

Query Match 72.1%; Score 31; DB 1; Length 594;
Best Local Similarity 55.6%; Pred. No. 93;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 NMLSEVERE 9
1:1:1111
Db 100 NMLSEVERE 108

RESULT 10
MUTL_HAEIN
ID MUTL_HAEIN STANDARD: PRT; 629 AA.
AC P44494;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE DNA mismatch repair protein mult.
GN MUTL OR H10067.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN=Rd / KM20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleschmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uterback T.R., Hanna M.C., Nguyen D.T., Sauder D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geophagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
RT Rd.";
RL Science 269:496-512(1995).
CC -I- FUNCTION: THIS PROTEIN IS INVOLVED IN THE REPAIR OF MISMATCHES IN
CC DNA. IT IS REQUIRED FOR DAM-DEPENDENT METHYL-DIRECTED DNA MISMATCH
CC REPAIR. MAY ACT AS A MOLECULAR MATCHMAKER, A PROTEIN THAT
CC PROMOTES THE FORMATION OF A STABLE COMPLEX BETWEEN TWO OR MORE
CC DNA-BINDING PROTEINS IN AN ATP-DEPENDENT MANNER WITHOUT ITSELF
CC BEING PART OF A FINAL EFFECTOR COMPLEX (BY SIMILARITY).
CC -I- SIMILARITY: BELONGS TO THE DNA MISMATCH REPAIR MUTL/HEXB FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC
CC -----
DR EMBL: U32692; AAC21745.1; -;
DR HSSP: P23367; IBKN.
DR TIGR: H10067; -;
DR InterPro: IPR003594; Amphib_Atpase.
DR InterPro: IPR002099; DNA_mis_repair.
DR Pfam: PF01119; DNA_mis_repair; 1.
DR Pfam: PF02518; HATPase_C; 1.
DR TIGRFAMs: TIGR00585; mult; 1.
DR PROSITE: PS00058; DNA_MISMATCH_REPAIR_1; 1.
KW DNA repair; Complete proteome.
SO SEQUENCE 629 AA; 71622 MW; 0ADC240FD94D1556 CRC64;

Query Match 72.1%; Score 31; DB 1; Length 629;
Best Local Similarity 75.0%; Pred. No. 99;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NMLSEVERE 8
1:1:1111
Db 596 NMLSEVERE 603

RESULT 11
RRF_HELPJ
ID RRF_HELPJ STANDARD: PRT; 185 AA.
AC Q925X1;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ribosome recycling factor (ribosome releasing factor) (RRF).
GN RRF OR JHP1177.
OS Helicobacter pylori J99 (Campylobacter pylori J99).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=85963;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99120557; PubMed=9923682;
RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
RA Trust T.J.;

RT "Genomic sequence comparison of two unrelated isolates of the human
RT gastric pathogen *Helicobacter pylori*."
RL Nature 397:176-180(1999).
CC -1- FUNCTION: RESPONSIBLE FOR THE RELEASE OF RIBOSOMES FROM MESSENGER
CC RNA AT THE TERMINATION OF PROTEIN BIOSYNTHESIS. MAY INCREASE THE
CC EFFICIENCY OF TRANSLATION BY RECYCLING RIBOSOMES FROM ONE ROUND OF
CC TRANSLATION TO ANOTHER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE RRF FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AEO01545; AAD06763.1; -.
DR HSSP: Q9X1B9; 1DD5.
DR InterPro: IPR002661; RRF.
DR Pfam: PF01765; RRF.1.
DR TIGRPFAMs: TIGR00496; frr.1.
KW Protein biosynthesis; Complete proteome.
SQ SEQUENCE 185 AA; 20917 MW; 34F9C0692C8CDB8 CRC64;

Query Match 69.8%; Score 30; DB 1; Length 185;
Best Local Similarity 62.5%; Pred. No. 44;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 NMLEVER 8
Db 73 NMLEVER 80
ID 1:1:1:1
ID RRF_HELPY STANDARD; PRT; 185 AA.
AC P56398;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ribosome recycling factor (Ribosome releasing factor) (RRF).
GN FRR OR HP1356.
OS *Helicobacter pylori* (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=26695 / ATCC 700392;
RX MEDLINE=97394467; PubMed=9252185;
RA Tomb J.-F., White O., Kervatage A.R., Clayton R.A., Sutton G.G.,
RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
RA Loftus B., Richardson D., Dodson R., Khalak H.G., Golder A.,
RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA Berg D.E., Gocayne J.D., Uterback T.R., Peterson J.D., Kelley J.M.,
RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Wathey L., Wallin E.,
RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA Venter J.C.;
RT "The complete genome sequence of the gastric pathogen *Helicobacter*
RT *pylori*."
RL Nature 388:539-547(1997).
CC -1- FUNCTION: RESPONSIBLE FOR THE RELEASE OF RIBOSOMES FROM MESSENGER
CC RNA AT THE TERMINATION OF PROTEIN BIOSYNTHESIS. MAY INCREASE THE
CC EFFICIENCY OF TRANSLATION BY RECYCLING RIBOSOMES FROM ONE ROUND OF
CC TRANSLATION TO ANOTHER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE RRF FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AEO00631; AAD08302.1; -.
DR HSSP: Q9X1B9; 1DD5.
DR TIGR: HP1256; -.
DR InterPro: IPR002661; RRF.
DR Pfam: PF01765; RRF.1.
DR TIGRPFAMs: TIGR00496; frr.1.
KW Protein biosynthesis; Complete proteome.
SQ SEQUENCE 185 AA; 20915 MW; 126AA4597C8CDBFE CRC64;

Query Match 69.8%; Score 30; DB 1; Length 185;
Best Local Similarity 62.5%; Pred. No. 44;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 NMLEVER 8
Db 73 NMLEVER 80
ID 1:1:1:1
ID VARD_METMA STANDARD; PRT; 209 AA.
AC 060188;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE V-type ATP synthase subunit D (EC 3.6.3.14) (V-type ATPase subunit D).
GN ATPD OR AHAD OR MM0778.
OS Methanosarcina mazei (Methanosarcina frisia).
OC Archaea; Euryarchaeota; Methanococci; Methanosarcinales;
OC Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2209;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Goel / GOI / ATCC BAA-199 / DSM 3647 / OCM 88;
RX MEDLINE=22120827; PubMed=12125824;
RA Deppe-Meier U., Johann A., Hartsch T., Merkl R., Schmitz R.A.,
RA Martinez-Arias R., Henne A., Wietzer A., Baumer S., Jacobl C.,
RA Brueggemann H., Lienard T., Christmann A., Boemcke M., Steckel S.,
RA Bhattacharyya A., Lykidis A., Overbeek R., Klenk H.-P., Gunsalus R.P.,
RA Fritz H.-J., Gottschalk G.;
RT "The genome of *Methanosarcina mazei*: evidence for lateral gene
RT transfer between Bacteria and Archaea."
RL J. Mol. Microbiol. Biotechnol. 4:453-461(2002).
CC -1- FUNCTION: PRODUCES ATP FROM ADP IN THE PRESENCE OF A PROTON
CC GRADIENT ACROSS THE MEMBRANE.
CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (in) = ADP + phosphate +
CC H(+) (out).
CC -1- SUBUNIT: IS COMPOSED OF SIX SUBUNITS: SUBUNITS A, B, C, D, E, F,
CC G.
CC -1- SIMILARITY: BELONGS TO THE V-ATPASE D SUBUNIT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL: U47274; AAC06377.1; -
 DR EMBL: AE013303; AAM30474.1; ALT_INIT.
 DR InterPro: IPR002699; ATPsynth_Dsub.
 DR Pfam: PF01813; ATP-synt_D; 1.
 DR Prodom: PD004122; ATPsynth_Dsub; 1.
 DR TIGRfam: TIGR00309; V_ATPase_subd; 1.
 DR KEGG: ATP synthetase; Hydrogen ion transport.
 SK SEQUENCE 209 AA; 23901 MW; 57025350BE07CDB CRC64;

Query Match 69.8%; Score 30; DB 1; Length 209;
 Best Local Similarity 75.0%; Pred. No. 50;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 NMLSEVERE 9
 DB 188 NMLEMERE 195

RESULT 14

DGK2_LACAC STANDARD; PRT; 223 AA.
 AC 059484;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Deoxyguanosine kinase (EC 2.7.1.113) (DGOU kinase) (DGK)
 OS (Deoxyguanoside kinase complex I F-component).
 OC Lactobacillus acidophilus.
 OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
 OC Lactobacillus.
 NC NCB1_TaxID=1579;

RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 11506 / R-26;
 RX MEDLINE=95204449; PubMed=7896798;
 RA Ma G.T., Hong Y.S., Ives D.H.;
 RT "Cloning and expression of the heterodimeric deoxyguanosine
 RL kinase/deoxyadenosine kinase of Lactobacillus acidophilus R-26."
 J. Biol. Chem. 270:6595-6601(1995).
 [2]

RP SEQUENCE OF 1-17.
 RC STRAIN-ATCC 11506 / R-26;
 RX MEDLINE=94227067; PubMed=8172906;
 RA Ikeda S., Ma G.T., Ives D.H.;

RT "Heterodimeric deoxyguanoside kinases of Lactobacillus acidophilus
 RL R-26: functional assignment of subunits using limited proteolysis
 controlled by end-product inhibitors.";
 Biochemistry 33:5328-5334(1994).

CC -1- FUNCTION: DGK/DAG PLAYS AN ESSENTIAL ROLE IN GENERATING THE
 CC DEOXYRIBONUCLEOTIDE PRECURSORS, DCTP AND DCTP, FOR DNA METABOLISM.
 CC -1- CATALYTIC ACTIVITY: ATP + deoxyguanosine = ADP + dGMP.
 CC -1- SUBUNIT: HETERODIMER OF A DEOXYADENOSINE (DAG) AND A
 CC DEOXYGUANOSINE KINASE (DGK).

CC -1- SIMILARITY: BELONGS TO THE DGK/DAG FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: U01881; AAB09751.1; -
 DR InterPro: IPR002624; dmk.
 DR Pfam: PF01712; dmk; 1.
 KW Transferase; Kinase; ATP-binding.

FT INIT_MET 0
 NP_BIND 7 14 ATP (POTENTIAL).
 SO SEQUENCE 223 AA; 26183 MW; 87EF74968751ED5D CRC64;

Query Match 69.8%; Score 30; DB 1; Length 223;
 Best Local Similarity 62.5%; Pred. No. 54;

Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NMLSEVER 8
 DB 110 NMLEMERE 117

RESULT 15

PT16_BOVIN STANDARD; PRT; 378 AA.
 ID PT16_BOVIN
 AC 002739;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Serine proteinase inhibitor B-43.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 NC NCB1_TaxID=9913;
 OX [1]

RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=97165893; PubMed=9013786;
 RA Nakaya N., Nishibori M., Kawabata M., Saeki K.;
 RT "Cloning of a serine proteinase inhibitor from bovine brain:
 RT expression in the brain and characterization of its target
 RT proteinases.";
 RL Brain Res. Mol. Brain Res. 42:293-300(1996).
 CC -1- FUNCTION: MAY BE INVOLVED IN THE REGULATION OF SERINE PROTEINASES
 CC PRESENT IN THE BRAIN OR EXTRAVASATED FROM THE BLOOD.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- TISSUE SPECIFICITY: BRAIN.
 CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: D55670; BAA19875.1; -
 DR HSSP: P05120; IBY7.
 DR InterPro: IPR000215; Serpin.
 DR Pfam: PF000079; serpin; 1.
 DR SMART: SM00093; SERPIN; 1.
 DR PROSITE: PS00284; SERPIN; 1.
 KW Serpin; Serine protease inhibitor.
 FT ACT_SITE 343 344 REACTIVE BOND (BY SIMILARITY).
 SO SEQUENCE 378 AA; 42560 MW; 664F499CCE263A CRC64;

Query Match 69.8%; Score 30; DB 1; Length 378;
 Best Local Similarity 75.0%; Pred. No. 93;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NMLSEVER 8
 DB 75 NMLEMERE 82

Search completed: November 18, 2002, 15:53:47
 Job time : 7.80645 secs

THIS PAGE BLANK (USPTO)